pK$_a$ measurements for the SAMPL6 prediction challenge for a set of kinase inhibitor-like fragments

Mehtap Işık[1,2]**, Dorothy Levorse**[3]**, Ariën S. Rustenburg**[1,4]**, Ikenna E. Ndukwe**[5]**, Heather Wang**[6]**, Xiao Wang**[5]**, Mikhail Reibarkh**[5]**, Gary E. Martin**[5]**, Alexey A. Makarov**[6]**, David L. Mobley**[7]**, Timothy Rhodes**[3]**, John D. Chodera**[1]*

[1]Computational and Systems Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY 10065, United States; [2]Tri-Institutional PhD Program in Chemical Biology, Weill Cornell Graduate School of Medical Sciences, Cornell University, New York, NY 10065, United States; [3]Pharmaceutical Sciences, MRL, Merck & Co., Inc., 126 East Lincoln Avenue, Rahway, New Jersey 07065, United States; [4]Graduate Program in Physiology, Biophysics, and Systems Biology, Weill Cornell Medical College, New York, NY 10065, United States; [5]Process and Analytical Research and Development, Merck & Co., Inc., Rahway, NJ 07065, United States; [6]Analytical Research & Development, MRL, Merck & Co., Inc., MRL, 126 East Lincoln Avenue, Rahway, New Jersey 07065, United States; [7]Department of Pharmaceutical Sciences and Department of Chemistry, University of California, Irvine, Irvine, California 92697, United States

*For correspondence:
john.chodera@choderalab.org (JDC)

Abstract Determining the net charge and protonation states populated by a small molecule in an environment of interest—such as solvent, a protein binding site, or a lipid bilayer—or the cost of altering those protonation states upon transfer to another environment is a prerequisite for predicting its physicochemical and pharmaceutical properties, as well as interactions with biological macromolecules using computational models. Incorrectly modeling the dominant protonation state, shifts in dominant protonation state, or the population of significant mixtures of protonation states can lead to large modeling errors that degrade the accuracy of physical modeling and hinder the ability to use physical modeling approaches for molecular design. For small molecules, the acid dissociation constant (pK$_a$) is the primary quantity needed to determine the ionic states populated by a molecule in an aqueous solution at a given pH. As a part of SAMPL6 community challenge, we organized a blind pK$_a$ prediction component to assess the accuracy with which contemporary pK$_a$ prediction methods can predict this quantity, with the ultimate aim of assessing the expected impact on modeling errors this would induce. While a multitude of approaches for predicting pK$_a$ values currently exist, predicting the pK$_a$s of drug-like molecules can be difficult due to challenging properties such as multiple titratable sites, heterocycles, and tautomerization. For this challenge, we focused on set of 24 small molecules selected to resemble selective kinase inhibitors—an important class of therapeutics replete with titratable moieties. Using a Sirius T3 instrument that performs automated acid-base titrations, we used UV-absorbance based pK$_a$ measurements to construct a high-quality experimental reference dataset of macroscopic pK$_a$s for the evaluation of computational pKa prediction methodologies that was utilized in the SAMPL6 pK$_a$ challenge. For several compounds in which the microscopic protonation states associated with macroscopic pK$_a$s were ambiguous, we performed follow-up NMR experiments to disambiguate the microstates involved in the transition. This dataset provides a useful standard benchmark dataset for the evaluation of pK$_a$ prediction methodologies on kinase inhibitor-like compounds.

42

Keywords

acid dissociation constants · spectrophotometric pK$_a$measurement · blind prediction challenge · SAMPL ·
macroscopic pK$_a$· microscopic pK$_a$· macroscopic protonation state · microscopic protonation state

Abbreviations

SAMPL Statistical Assessment of the Modeling of Proteins and Ligands

pK$_a$ $-log_{10}$ acid dissociation equilibrium constant

p$_s$K$_a$ $-log_{10}$ apparent acid dissociation equilibrium constant in cosolvent

DMSO Dimethyl sulfoxide

ISA Ionic-strength adjusted

SEM Standard error of the mean

TFA Target factor analysis

LC-MS Liquid chromatography - mass spectrometry

NMR Nuclear magnetic resonance spectroscopy

HMBC Heteronuclear Multiple-Bond Correlation

TFA-*d* deutero-trifluoroacetic acid

Introduction

SAMPL (Statistical Assessment of the Modeling of Proteins and Ligands) is a recurring series of blind prediction
challenges for the computational chemistry community [1, 2]. Through these challenges, SAMPL aims to
evaluate and advance computational tools for rational drug design. By focusing the community on specific
phenomena relevant to drug discovery poorly predicted by current models, isolating that phenomenon
from other confounding factors in well-designed test systems, evaluating tools prospectively, enforcing
data sharing to learn from failures, and releasing the resulting high-quality datasets into the community as
benchmark sets, SAMPL has driven progress in a number of areas over seven previous rounds of challenge
cycles [3–7, 7–15].

As a stepping stone to enabling the accurate prediction of protein-ligand binding affinities, SAMPL
has focused on evaluating how well physical and empirical modeling methodologies can predict various
physicochemical properties relevant to binding and drug discovery, such as hydration free energies (which
model aspects of desolvation in isolation), distribution coefficients (which model transfer from aqueous to
nonpolar environments without conflating slow protein dynamics), and host-guest binding affinities (which
model high-affinity association with fewer degrees if freedom and without the complication of slow protein
dynamics).

These physicochemical property prediction challenges—in addition to assessing predictive accuracy of
quantities that are useful in various stages of drug discovery in their own right—have been helpful in pinpoint
deficiencies in computational models that can lead to substantial errors in affinity predictions. Errors is
pK$_a$predictions can cause modeling the wrong charge, protonation and tautomerization states which effects
hydrogen bonding opportunities and overall dipole moment of the ligand.

JDC: Cite some SAMPL papers highlighting lessons learned.

Neglect of protonation state effects can lead to large modeling errors

As part of the SAMPL5 challenge series, a new cyclohexane-water distribution constant (logD) prediction
challenge was introduced, where participants predicted the transfer free energy of small drug-like molecules
between an aqueous buffer phase at pH 7.4 and a nonaqueous cyclohexane phase [16, 17]. While octanol-
water distribution coefficient measurements are more common, cyclohexane was selected for the simplicity
of its liquid phase and relative dryness compared to wet octanol phases. While the expectation was that this
challenge would be relatively straightforward given the lack of complexity of cyclohexane phases, analysis of
participant performance revealed that multiple factors contributed to significant prediction failures: poor

88 conformational sampling of flexible solute molecules, misprediction of relevant protonation and tautomeric
89 states (or failure to accommodate shifts in their populations), and force field inaccuracies resulting in
90 bias towards the cyclohexane phase. While these findings justified the benefit of future iterations of blind
91 distribution or partition coefficient challenges, the most surprising observation from this initial logD challenge
92 was that participants almost uniformly neglected to accurately model protonation state effects, and that
93 neglect of these effects led to surprisingly large errors in transfer free energies [16–18]. Careful quantum
94 chemical assessments of the magnitude of these protonation state effects found that their neglect could
95 introduce errors up to 6–8 kcal/mol for some compounds [18]. This effect stems from the need to account
96 for the free energy difference between the major ionization state in cyclohexane (most likely neutral state)
97 and in water phase (which could be neutral or charged). To isolate this surprisingly large protonation state
98 effects modeling error from difficulties related to lipophilicity (logP and logD) prediction methods, we decided
99 to organize a set of staged physicochemical property challenges using a consistent set of molecules that
100 resemble small molecule kinase inhibitors—an important drug class replete with multiple titratable moieties.
101 This series of challenges will first evaluate the ability of current-generation modeling tools to predict acid
102 dissociation constants (pK_a). It will be followed by a partition/distribution coefficient challenge to evaluate
103 the ability to incorporate experimentally-provided pK_a values into prediction of distribution coefficients to
104 ensure methodologies can correctly incorporate protonation state effects into their predictions. A third
105 stage of challenge will follow: a new blinded partition/distribution coefficient challenge where participants
106 must predict pK_a values on their own. At the conclusion of this series of challenges, we will have ensured that
107 modern physical and empirical modeling methods have eliminated this large source of spurious errors from
108 modeling both simple and complex phenomena.

109   This article reports on the experiments for the first stage of this series of challenges: SAMPL6 pK_a prediction
110 challenge. The selection of a small molecule set and collection of experimental pK_a data are described in
111 detail.

Conceptualization of a blind pK_a challenge

113 This is the first time a blind pK_a prediction challenge has been fielded as part of SAMPL. In this first iteration of
114 the challenge, we aimed to assess the performance of current pK_a prediction methods and isolate potential
115 causes of inaccurate pK_a estimates.

116   The prediction of pK_a values for drug-like molecules can be complicated by several effects: the presence
117 of multiple (potentially coupled) titratable sites, the presence of heterocycles, tautomerization, the confor-
118 mational flexibility of large molecules, and ability of intramolecular hydrogen bonds to form. We decided
119 to focus on the chemical space of small molecule kinase inhibitors in the first iteration of pK_a prediction
120 challenge. A total of 24 small organic molecules (17 fragment-like and 7 drug-like) were selected for their
121 similarity to known small molecule kinase inhibitors, while also considering properties predicted to affect the
122 experimental tractability of pK_a and logP measurements such as solubility and predicted pK_as. Macroscopic
123 pK_a values were collected experimentally with UV-absorbance spectroscopy based pK_a measurements using
124 a Sirius T3 instrument, which automates the sample handling, titration, and spectroscopic measurement
125 process to allow high-quality measurements. Sirius T3 is equiped with an autosampler which allowed us to
126 run 8-10 measurements per day. Experimental data was kept blinded for three months (25 Oct 2017 through
127 23 Jan 2018) to allow participants in the SAMPL6 pK_a challenge to submit truly blinded computational predic-
128 tions. Eleven research groups participated in this challenge, providing a total of 93 prediction submission
129 sets that cover a large variety of contemporary pK_a prediction methods.

Our selected experimental approach determines macroscopic pK_a values

131 Whenever experimental pK_a measurements are used for evaluating pK_a predictions, it is important to differen-
132 tiate between microscopic and macroscopic pK_a values. In molecules containing multiple titratable moieties,
133 the protonation state of one group can affect the proton dissociation propensity of another functional group.
134 In such cases, the **microscopic pK_a** (group pK_a) refers to the pK_a of deprotonation of a single titratable group
135 while all the other titratable and tautomerizable functional groups of the same molecule are held fixed.
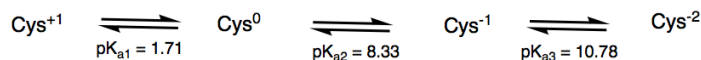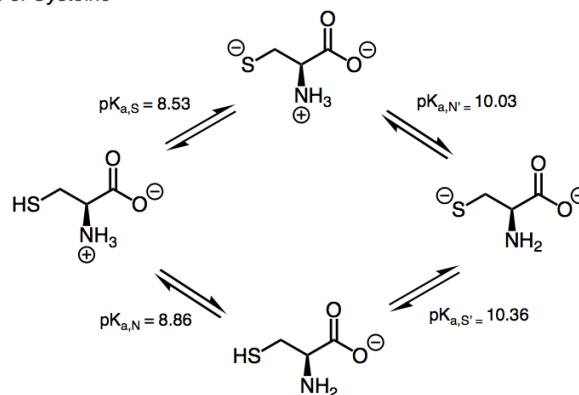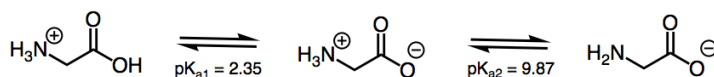136 Different protonation states and tautomer combinations constitute different microstates. The **macroscopic**

Figure 1. Assignment of cysteine and glycine pK$_a$ values. pK_{a1}, pK_{a2}, and pK_{a3} are macroscopic acid dissociation constants for cysteine and glycine [24]. When pK$_a$ values of a polyprotic molecule are very different, such as in the case of glycine, it is possible to assign the pK$_a$s to individual groups since the dissociation of protons is stepwise [19]. However, stepwise dissociation can not be assumed for cysteine, because pK_{a2} and pK_{a3} are very close in value. Four underlying microscopic pK$_a$s ($pK_{a,S}$, $pK_{a,N}$, $pK_{a,S'}$, and $pK_{a,N'}$) for cysteine were measured using UV spectra analysis of cysteine and derivatives [25]. Notice that the closeness of $pK_{a,S}$ and $pK_{a,N}$ values indicates similar probability of proton dissociation from these groups. This figure is adopted from [19].

---

137   **pK$_a$** (molecular pK$_a$) defines the acid dissociation constant related to the observable loss of a proton from
138   a molecule regardless of which functional group the proton is dissociating from, so it doesn't necessarily
139   convey structural information.

140   Whether a measured pK$_a$ is microscopic or macroscopic depends on the experimental method used
141   (Figure 2). For a molecule with only one titratable proton, the microscopic pK$_a$ is equal to the macroscopic
142   pK$_a$. For a molecule with multiple titratable groups, however, throughout a titration from acidic to basic
143   pH, the deprotonation of some functional groups can take place almost simultaneously. For these multi-
144   protic molecules, the experimentally-measured macroscopic pK$_a$ will include contributions from multiple
145   microscopic pK$_a$s with similar values (i.e., acid dissociation of multiple microstates). Cysteine provides an
146   example of this behavior with its two macroscopic pK$_a$s observable by spectrophotometric or potentiometric
147   pK$_a$ measurement experiments [19, 20].

148   While four microscopic pK$_a$s can be defined for cysteine, experimentally observed pK$_a$ values can't be
149   assigned to individual functional groups directly (Figure 1), and more advanced techniques capable of
150   resolving individual protonation sites—such as NMR [21], Raman spectroscopy [22, 23], and analysis of pK$_a$s
151   in molecular fragments or derivatives—are required to unambiguously assign the site of protonation state
152   changes. On the other hand, when there is a large difference between microscopic pK$_a$s in a multiprotic
153   molecule, the proton dissociations won't overlap and macroscopic pK$_a$s observed by experiments can
154   be assigned to individual titratable groups. The pK$_a$ values of glycine provide a good example of this
155   scenario [19, 20, 22]. We recommend a short review on the assignment of pK$_a$ values authored by Ivan G.
156   Darvey [20] for a good introduction to the concepts of macroscopic *vs* microscopic pK$_a$s.

157   The most common methods for measuring small molecule pK$_a$ values are UV-absorbance spectroscopy
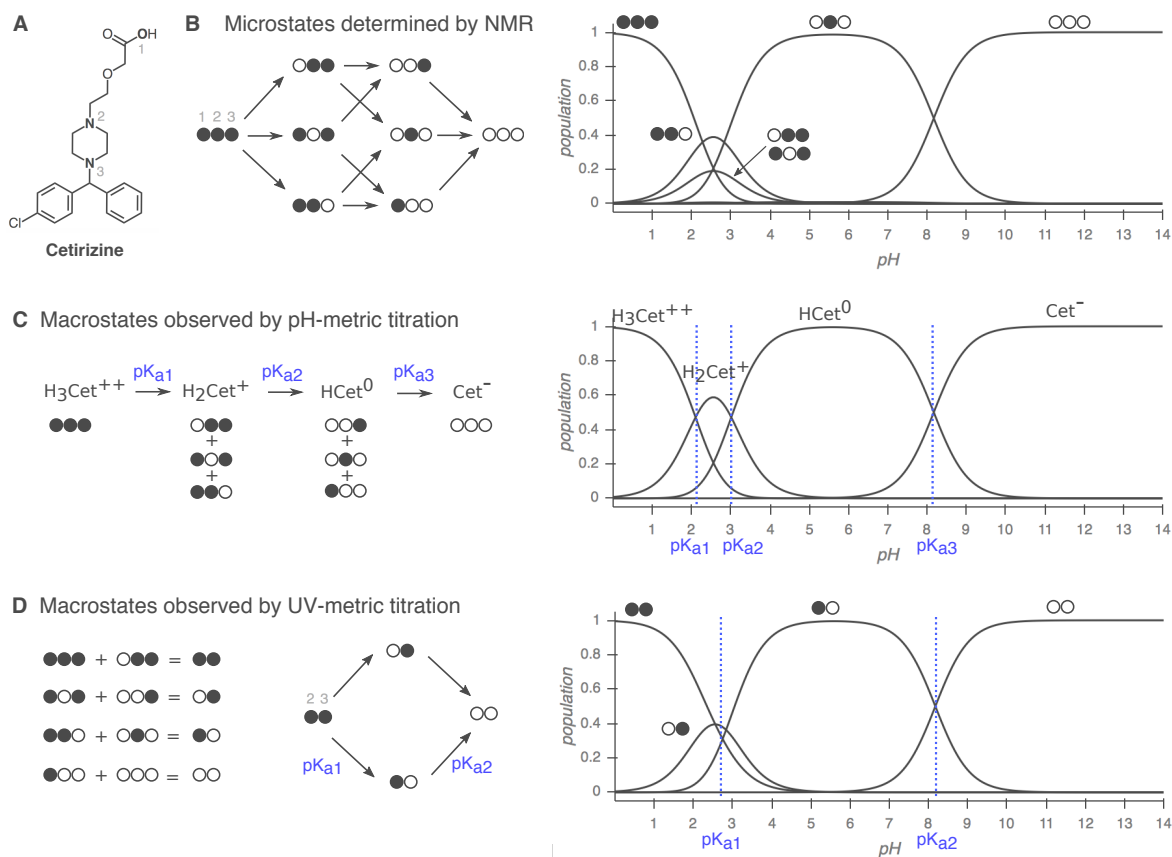
Figure 2. Comparison of macroscopic and microscopic pK$_a$ measurement methods. Filled circles represent protonated sites and empty circles represent deprotonated sites with the order of carboxylic acid (1), piperazine nitrogen (2), and piperazine nitrogen (3). Protonation state populations shown for pH-metric and UV-metric pK$_a$ measurement methods are simulations, calculated using NMR-based microscopic pK$_a$ values. (**A**) Cetirizine has $n = 3$ titratable sites, shown in bold. (**B**) *Left:* The 8 microstates (2^n) and 12 microscopic pK$_a$s ($n2^{n-1}$) of cetirizine. *Right:* Relative population of microspecies with respect to pH. Potentially all microstates can be resolved via NMR. (**C**) Simulated pH-metric (potentiometric) titration and macroscopic populations. For a polyprotic molecule, only macroscopic pK$_a$s can be measured with pH-metric titration. Microstates with different total charge (related to the number of protons) can be resolved, but microstates with the same total charge are observed as one macroscopic population. (**D**) Simulated microscopic populations for UV-metric (spectrophotometric) titration of cetirizine. Since only protonation of the titration sites within four heavy atoms of the UV-chromophore is likely to cause an observable change in the UV-absorbance spectra, microstates that only differ by protonation of the distal carboxylic acid cannot be differentiated. Moreover, populations that overlap may or may not be resolvable depending on how much their absorbance spectra in the UV region differ. Both UV-metric and pH-metric pK$_a$ determination methods measure macroscopic pK$_a$s for polyprotic molecules, which cannot easily be assigned to individual titration sites and underlying microstate populations in the absence of other experimental evidence that provides structural resolution, such as NMR. Note that macroscopic populations observed in these two methods are composed of different combinations of microstates depending on the principles of measurement technique. Here, the illustrative diagram style was adopted from [26], and NMR-determined microscopic pK$_a$s for cetirizine were taken from [27].

158 (UV-metric titration) [28–30], potentiometry (pH-metric titration) [30, 31], capillary electrophoresis [32, 33],
159 and NMR spectroscopy [21] with NMR being the most expensive method. Other, less popular pK$_a$ measurement
160 techniques include conductometry, HPLC, solubility or partition based estimations, calorimetry, fluorome-
161 try, and polarimetry [34]. UV-metric and pH-metric methods (Figure 3) are limited to measuring aqueous
162 pK$_a$ values between 2 and 12 due to limitations of the pH-electrode used in these measurements. The
163 pH-metric method relies on determining the stoichiometry of bound protons with respect to pH, calculated
164 from volumetric titration with acid or base solutions. Accurate pH-metric measurements require high
165 concentrations of analyte as well as analytically prepared acid/base stocks and analyte solutions. By contrast,
166 UV-metric pK$_a$ measurements rely on the differences in UV-absorbance spectra of different protonation
167 states, generally permitting lower concentrations of analyte to be used. The pH and UV absorbance of the
168 analyte solution are monitored during titration.

169 Both UV-metric and pH-metric pK$_a$ determination methods measure macroscopic pK$_a$s for polyprotic
170 molecules, which can not be easily assigned to individual titration sites and underlying microstate popu-
171 lations in the absence of other experimental evidence that provides structural information, such as NMR
172 (Figure 2). Macroscopic populations observed in these two methods are composed of different combinations
173 of microstates depending on the principles of measurement technique. In potentiometric titrations, mi-
174 crostates with same total charge will be observed as one macrostate, while in spectrophotometric titrations,
175 protonation sites remote from chromophores might be spectroscopically invisible, and macrostates will be
176 formed from collections of microstates that manifest similar UV-absorbance spectra.

177 Spectrophotometric pK$_a$ determination is more sensitive than potentiometric determination, requiring
178 low analyte concentrations (50–100 μM)—especially advantageous for compounds with low solubilities—
179 but is only applicable to titration sites near chromophores. For protonation state changes to affect UV
180 absorbance, a useful rule of thumb is that the protonation site should be a maximum of four heavy atoms
181 away from the chromophore, which might consist of conjugated double bonds, carbonyl groups, aromatic
182 rings, etc. Although potentiometric measurements do not suffer from the same observability limitations,
183 higher analyte concentrations (∼5 mM) are necessary for the analyte to provide sufficiently large enough
184 buffering capacity signal above water to produce an accurate measurement. The accuracy of pK$_a$s fit to
185 potentiometric titrations can also be also sensitive to errors in the estimated concentration of the analyte in
186 the sample solution, while UV-metric titrations are insensitive to concentration errors. We therefore decided
187 to adopt spectrophotometric measurements for collecting the experimental pK$_a$ data for this challenge, and
188 selected a compound set to ensure that all potential titration sites are in the vicinity of UV-chromophores.

189 Here, we report on the selection of SAMPL6 pK$_a$ challenge compounds, their macroscopic pK$_a$ values
190 measured by UV-metric titrations using a Sirius T3, and NMR-based microstate characterization of two
191 SAMPL6 compounds with ambiguous protonation states associated with the observed macroscopic pK$_a$s
192 (SM07 and SM14). We discuss implications of the use of this experimental technique for the interpretation of
193 pK$_a$ data, and provide suggestions for future pK$_a$ data collection efforts with the goal of evaluating or training
194 computational pK$_a$ predictions.

Methods

Compound selection and procurement

197 To select a set of small molecules focusing on the chemical space representative of kinase inhibitors for
198 physicochemical property prediction challenges (pK$_a$ and lipophilicity) we started from the kinase-targeted
199 subclass of the ZINC15 chemical library [35] and applied a series of filtering and selection rules as depicted
200 in Figure 4A. We focused on the availability "now" and reactivity "anodyne" subsets of ZINC15 in the first
201 filtering step [http://zinc15.docking.org/subclasses/kinase/substances/subsets/now+anodyne/]. The "now"
202 label indicates the compounds were availabile for immediate delivery, while the "anodyne" label excludes
203 compounds matching filters that flag compounds with the potential for reactivity or pan-assay interference
204 (PAINs) [36, 37].

205 Next, we identified resulting molecules that were also available for procurement through eMolecules [38]
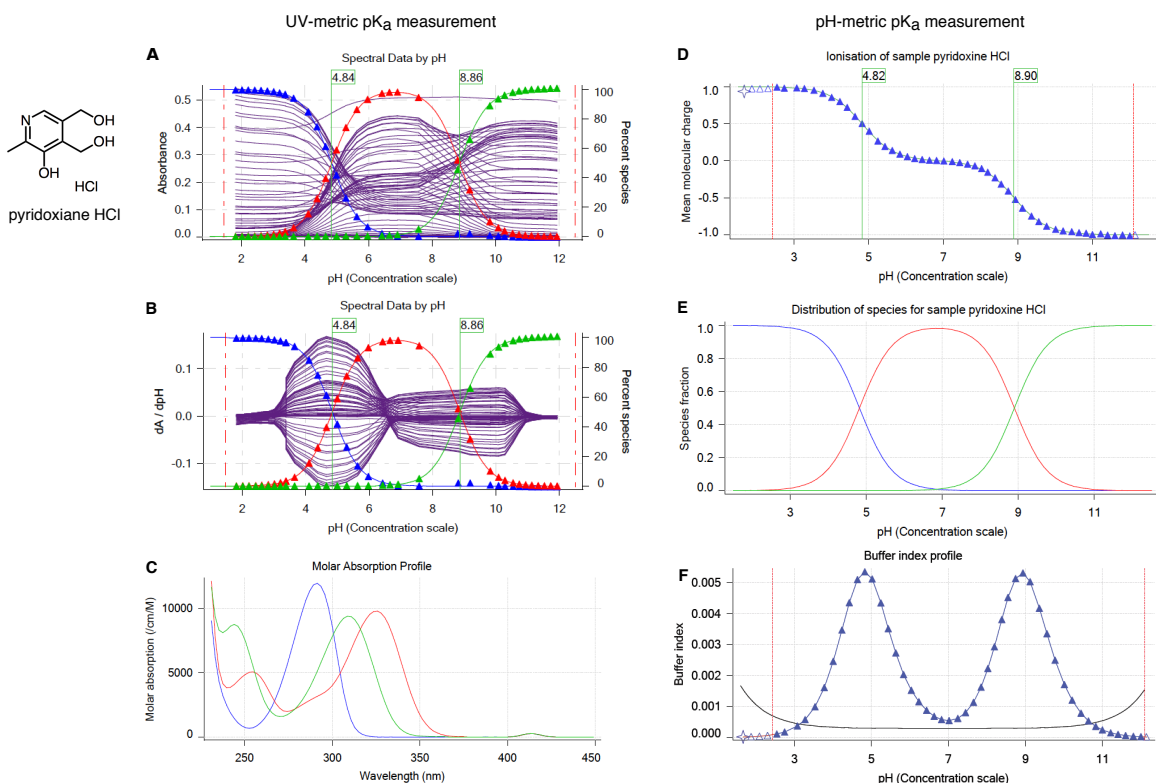206 (free version, downloaded 1 June 2017), the supplier that would be used for procurement in this exercise. To

Figure 3. UV-metric (spectrophotometric) and pH-metric (potentiometric) pK$_a$ measurements of pyridoxine HCl with Sirius T3. Spectrophotometic pK$_a$ measurement (panels **A, B, C**) relies on differences in the UV absorbance spectra between microscopic protonation states to deconvolute the population of macrostate species as a function of pH. While highly sensitive (and therefore requiring a very low analyte concentration of ~ 50 μM), this approach can only resolve changes in protonation states for titratable sites near chromophores and cannot separate the populations of microstates that change in the same manner as a function of pH. (**A**) Multiwavelength UV absorbance vs pH. Purple lines represents absorbance at distinct wavelengths in UV region. (**B**) Derivative of multiwavelength absorbance with respect to pH (dA/dpH) vs pH is plotted with purple lines. In **A** and **B** Blue, red, and green triangles represent population of protonation states (from most protonated to least protonated) as calculated from a global fit to experimental UV absorbances for all pH values, while thin lines denote model fits that utilize the fitted model pK$_a$s to compute populations. pK$_a$ values (green flags) correspond to inflection point of multiwavelength absorbance data where change in absorbance with respect to pH is maximum. (**C**) Molar absorption coefficients vs wavelength for each protonation state as resolved by TFA. **D, E, F** illustrate potentiometric pK$_a$ measurement where molar addition of acid or base is tracked as pH is titrated. (**D**) Mean molecular charge vs pH. Mean molecular charge is calculated based on the model provided for the analyte: predicted number and nature of titratable sites (acid or base type), and number of counter ions present. pK$_a$ values are calculated as inflection points of charge vs pH plot. (**E**) Predicted macroscopic protonation state populations vs pH calculated based on pK$_a$ values (H$_2$A$^+$: blue, HA: red, and A$^-$: green) (**F**) Buffering index vs pH profile of water (grey solid line, theoretical) and the sample solution (blue triangles represent experimental data points). A higher concentration of analyte (~ 5 mM) is necessary for the potentiometric method than the spectrophotometric method in order to provide large enough buffering capacity signal above water for an accurate measurement.

207 find the intersection of ZINC15 kinase subset and eMolecules database, we matched molecules using their
208 canonical isomeric SMILES strings, as computed via the OpenEye OEChem Toolkit [39].

209   To extract availability and price information from eMolecules, we queried using a list of SMILES (as
210 reported in eMolecules database) of the intersection set. We further filtered the intersection set (1204
211 compounds) based on delivery time (Tier 1 suppliers, 2-week delivery) and at least 100 mg availability
212 in powder form (format: Supplier Standard Vial). We aimed to purchase 100 mg of each compound in
213 powder form with at least 90% purity. We calculated 100 mg was enough for optimization and replicate
214 experiments to measure pK_a, logP and solubility measurements with the Sirius T3. Each UV-metric and
215 pH-metric pK_a measurement requires a minimum of 0.01 mg and 1.00 mg compound (solid or delivered via
216 DMSO stock solution), respectively. logP and pH-dependent solubility measurements require around 2 mg
217 and 10 mg of solid chemical, respectively.

Filtering for predicted measurable pK_as and lack of experimental data

219 The Sirius T3 (Pion) instrument used to collect pK_a and logP/logD measurements requires a titratable group
220 in the pK_a range of 2–12, so we aimed to select compounds with predicted pK_as in the range of 3–11 to allow
221 a ~1 pK unit margin of error in pK_a predictions. pK_a predictions for compound selection were calculated using
222 Epik (Schödinger) sequential pK_a prediction (scan) [40, 41] with target pH 7.0 and tautomerization allowed
223 for generated states. We filtered out all compounds that did not have any pK_as predicted between 3–11, as
224 well as compounds with two predicted pK_a values predicted to be less than 1 pK_a unit apart in the hopes that
225 individual pK_as of multiprotic compounds could be resolved with spectrophotometric pK_a measurements.
226 With the goal of selecting compounds suitable for subsequent logP measurements, we eliminated com-
227 pounds with OpenEye XlogP [42] values less than -1 or greater than 6. Subsets of compounds with molecular
228 weights between 150–350 g/mol and 350–500 g/mol were selected for fragment-like and drug-like categories
229 respectively. Compounds without available price or stock quantity information were eliminated. As the goal
230 was to provide a blind challenge, compounds with publicly available experimental logP measurements were
231 also removed. The sources we checked for publicly available experimental logP values were the following:
232 DrugBank [43] (queried with eMolecules SMILES), ChemSpider [44] (queried by canonical isomeric SMILES),
233 NCI Open Database August 2006 release [45], Enhanced NCI Database Browser [46] (queried with canonical
234 isomeric SMILES), and PubChem [47] (queried with InChIKeys generated from canonical isomeric SMILES
235 with NCI CACTUS Chemical Identifier Resolver [48].)

Filtering for kinase inhibitor-like scaffolds

237 In order to include common scaffolds found in kinase inhibitors, we analyzed the frequency of rings found
238 in FDA-approved kinase inhibitors via Bemis-Murcko fragmentation [49, 50]. Heterocycles found more
239 than once in FDA-approved kinase inhibitors are shown in Figure 4B. In selecting 25 compounds for the
240 fragment-like set and 10 compounds for the drug-like set, we prioritized including at least one example of
241 each heterocycle, although we failed to find compounds with piperazine and indazole that satisfied all other
242 selection criteria. We observed that certain heterocycles (shown in Figure 4C) were overrepresented based
243 on our selection criteria; therefore, we limited the number of these structures in the SAMPL6 challenge set
244 to at most one in each set. To achieve broad and uniform sampling of the measurable logP dynamic range,
245 we segregated the molecules into bins of predicted XlogP values and selected compounds from each bin,
246 prioritizing less expensive compounds.

Filtering for UV chromophores

248 The presence of UV chromophores (absorbing in the 200–400 nm range) in close proximity to protonation
249 sites is necessary for spectrophotometric pK_a measurements. To filter for molecules with UV chromophores,
250 we looked at the substructure matches to the SMARTS pattern `[n,o,c][c,n,o]cc` which was considered
251 the smallest unit of pi-conjugation that can constitute a UV chromophore. This SMARTS pattern describes
252 extended conjugation systems comprised of four heavy atoms and composed of aromatic carbon, nitrogen,
253 or oxygen, such as 1.3-butadiene, which possesses an absorption peak at 217 nm. Additionally, the final set
254 of selected molecules was manually inspected to makes sure all potentially titratable groups were no more
255 than four heavy atoms away from a UV-chromophore.

25 fragment-like and 10 drug-like compounds were selected, out of which procurement of 28 was completed in time. pK_a measurements for 17 (SM01-SM17) and 7 (SM18-SM24) were successful, respectively. The resulting set of 24 small molecules constitute the SAMPL6 pK_a challenge set. For the other 4 compounds, UV-metric pK_a measurements show no detectable pK_as in the range of 2–12, so we decided not to include them in the SAMPL6 pK_a challenge. Experiments for these 4 compounds are not reported in this publication.

Python scripts used in the compound selection process are available from GitHub [https://github.com/choderalab/sampl6-physicochemical-properties]. Procurement details for each compound can be found in Supplementary Table 1. Chemical properties used in selection of compounds are summarized in Supplementary Table 2.

UV-metric pK_a measurements

Experimental pK_a measurements were collected using the spectrophotometric pKa measurement method with a Sirius T3 automated titrator instrument (Pion) at 25°C and constant ionic strength. The Sirius T3 is equipped with an Ag/AgCl double-junction reference electrode to monitor pH, a dip probe attached to UV spectrophotometer, a stirrer, and automated volumetric titration capability. The Sirius T3 UV-metric pK_a measurement protocol measures the change in multi-wavelength absorbance in the UV region of the absorbance spectrum while the pH is titrated over pH 1.8–12.2 [28, 29]. UV absorbance data is collected for 160–760 nm and 250–450 nm region is typically used for pK_a measurements. Subsequent global data analysis identifies pH-dependent populations of macrostates and fits one or more pK_a values to match this population with a pH-dependent model.

DMSO stock solutions of each compound with 10 mg/ml concentration were prepared by weighing 1 mg of powder chemical with a Sartorius Analytical Balance (Model: ME235P) and dissolving it in 100 μL DMSO (Dimethyl sulfoxide, Fisher Bioreagents, CAT: BP231-100, LOT: 116070, purity ≥ 99.7%). DMSO stock solutions were capped immediately to limit water absorption from the atmosphere due to the high hygroscopicity of DMSO and sonicated for 5–10 minutes in a water bath sonicator at room temperature to ensure proper dissolution. These DMSO stock solutions were stored in room temperature up to 2 weeks in capped glass vials. 10 mg/ml DMSO solutions were used as stock solutions for the preparation of three replicate samples for the independent titrations. For each experiment, 1–5 μL of 10 mg/ml DMSO stock solution was delivered to a 4 mL Sirius T3 glass sample vial with an electronic micropipette (Rainin EDP3 LTS 1-10 μL). The volume of delivered DMSO stock solution, which determines the sample concentration following dilution by the T3, is optimized individually for each compound to achieve sufficient but not saturated absorbance signal (targeting 0.5–1.0 AU) in the linear response region. Another limiting factor for sample concentration was ensuring that the compound remains soluble throughout the entire pH titration range. An aliquot of 25 μL of mid-range buffer (14.7 mM K_2HPO_4 and 0.15 M KCl in H_2O) was added to each sample, transfered with a micropipette (Rainin EDP3 LTS 10-100 μL) to provide enough buffering capacity in middle pH ranges so that pH could be controlled incrementally throughout the titration.

pH is temperature and ionic-strength dependent. A peltier device on the Sirius T3 kept the analyte solution at 25.0 ± 0.5 °C throughout the titration. Sample ionic strength was adjusted by dilution in 1.5 mL ionic strength-adjusted water (ISA water = 0.15 M KCl in H_2O) by the Sirius T3. Analyte dilution, mixing, acid/base titration, and measurement of UV-absorbance was automated by the Sirius T3 UV-metric pK_a measurement protocol. The pH was titrated between pH 1.8 and 12.2 via the addition of acid (0.5 M HCl) and base (0.5 M KOH), targeting 0.2 pH steps between UV absorbance spectrum measurements. Titrations were performed under argon flow on the surface of the sample solution to limit the absorption of carbon dioxide from air, which can alter the sample pH to a measurable degree. To fully capture all sources of experimental variability, instead of performing three sequential pH titrations in the same sample solution, three replicate samples (prepared from the same DMSO stock solution) were subjected to one round of pH titration each. Although this choice reduced throughput and increased analyte consumption, it limited the dilution of the analyte during multiple titrations, resulting in stronger absorbance signal for pK_a fitting. Under circumstances where analyte is scarce, it is also possible to do three sequential titrations using the same sample to limit consumption when the loss of accuracy is acceptable.

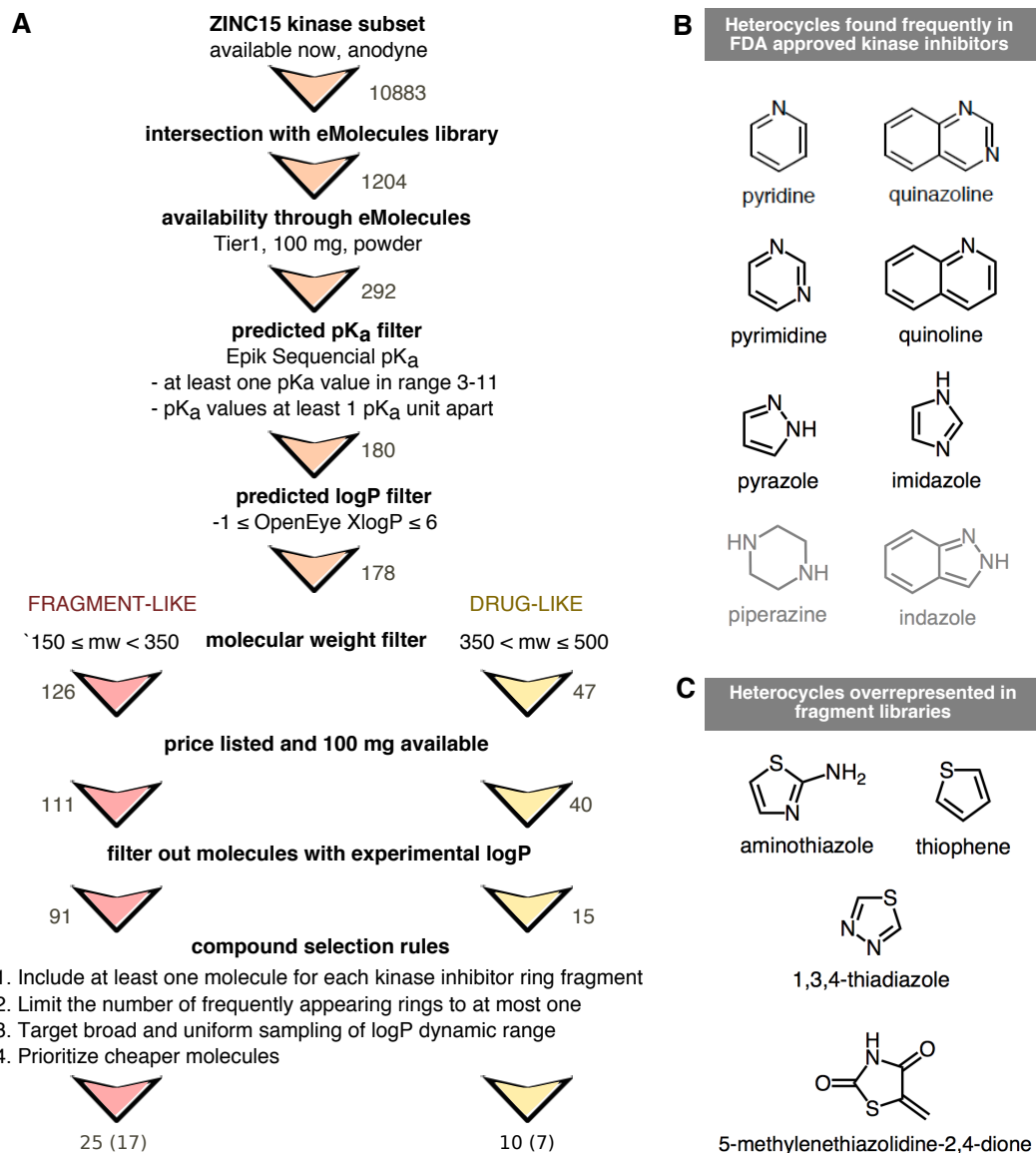Visual inspection of the sample solutions after titration and inspection of the pH-dependent absorbance

A

ZINC15 kinase subset
available now, anodyne

10883

intersection with eMolecules library

1204

availability through eMolecules
Tier1, 100 mg, powder

292

predicted pK$_a$ filter
Epik Sequencial pK$_a$
- at least one pKa value in range 3-11
- pK$_a$ values at least 1 pK$_a$ unit apart

180

predicted logP filter
$-1 \leq$ OpenEye XlogP ≤ 6

178

FRAGMENT-LIKE   **molecular weight filter**   DRUG-LIKE

`$150 \leq mw < 350$` $350 < mw \leq 500$

126   47

price listed and 100 mg available

111   40

filter out molecules with experimental logP

91   15

compound selection rules

1. Include at least one molecule for each kinase inhibitor ring fragment
2. Limit the number of frequently appearing rings to at most one
3. Target broad and uniform sampling of logP dynamic range
4. Prioritize cheaper molecules

25 (17)   10 (7)

B Heterocycles found frequently in FDA approved kinase inhibitors

pyridine   quinazoline

pyrimidine   quinoline

pyrazole   imidazole

piperazine   indazole

C Heterocycles overrepresented in fragment libraries

aminothiazole   thiophene

1,3,4-thiadiazole

5-methylenethiazolidine-2,4-dione

Figure 4. Compound selection for the SAMPL6 pK$_a$ challenge, with the goal of running subsequent logP/logD challenges on the same compound set. (**A**) Flowchart of filtering steps for the selection of compounds that resemble kinase inhibitors and their fragments. Numbers next to arrows indicate the number of compounds remaining after each filtering step. A total of 25 fragment-like and 10 drug-like compounds were selected, out of which procurement and pK$_a$ measurements for 17 fragment-like and 7 drug-like compounds were successful, respectively. (**B**) Frequent heterocycles found in FDA approved kinase inhibitors, as determined by Bemis-Murcko fragmentation into rings [49]. Black structures were represented in SAMPL6 set at least once. Compounds with piperazine and indazole (gray structures) could not be included in the challenge set due to library and selection limitations. (**C**) Structures of heterocycles that were overrepresented based on our compound selection workflow. We have limited occurance of these heterocycles to at most one.

shift in the 500–600 nm region of the UV spectra was used to verify no detectable precipitation occurred during the course of the measurement. Increased absorbance in the 500–600 nm region of the UV spectra is associated with scattering of longer wavelengths of light in the presence of colloidal aggregates. For each analyte, we optimized analyte concentration, direction of the titration, and pH titration range in order to maintain solubility over the entire experiment. The titration direction was specified so that each titration would start from the pH where the compound is most soluble: low-to-high pH for bases and high-to-low pH for acids. While UV-metric pK$_a$ measurements can be performed with analyte concentrations as low as 50 μM (although this depends on the absorbance properties of the analyte), some compounds may yet not be soluble at these low concentrations throughout the pH range of the titration. As the sample is titrated through a wide range of pH values, it is likely that low-solubility ionization states—such as neutral and zwitterionic states—will also be populated, limiting the highest analyte concentration that can be titrated without encountering solubility issues. For compounds with insufficient solubility to accurately determine a pK$_a$ value directly in a UV-metric titration, a cosolvent protocol was used, as described in the next section (**UV-metric pK$_a$ measurement with cosolvent**).

Two Sirius T3 computer programs—Sirius T3 Control v1.1.3.0 and Sirius T3 Refine v1.1.3.0—were used to execute measurement protocols and analyze pH-dependent multivelenghth spectra, respectively. Protonation state changes at titratable sites near chromophores will modulate the UV-absorbance spectra of these chromophores, allowing populations of distinct UV-active species to be resolved as a function of pH. To do this, basis spectra are identified and populations extracted via TFA analysis of the pH-dependent multi-wavelength absorbance [29]. When fitting the absorbance data to a titratable molecule model to estimate pK$_a$s, we selected the minimum number of pK$_a$s sufficient to provide a high-quality fit between experimental and modeled data based on visual inspection of pH-dependent populations.

This method is capable of measuring pK$_a$ values between 2–12 when protonatable groups are at most 4–5 heavy atoms away from chromophores such that a change in protonation state alters the absorbance spectrum of the chromophore. We selected compounds where titratable groups are close to potential chromophores (generally aromatic ring systems), but the possibility exists that our experiments did not detect protonation state changes of titratable groups distal from UV chromophores.

UV-metric pK$_a$ measurement with cosolvent

If analytes are not sufficiently soluble during the titration, pK$_a$ values cannot be accurately determined. If precipitation occurs, the UV-absorbance signal from pH-dependent precipitate formation can not be differentiated from the pH-dependent signal of soluble microstate species. For compounds with low aqueous solubility, pK$_a$ values were estimated from multiple apparent pK$_a$ measurements performed in ISA methanol:ISA water cosolvent solutions with various mole fractions, from which the pK$_a$ at 0% methanol (100% ISA water) can be extrapolated. This method is referred to as a UV-metric psKa measurement in the Sirius T3 Manual [51].

The cosolvent spectrophotometric pK$_a$ measurement protocol was very similar to the standard aqueous UV-metric pK$_a$ measurement protocol, with the following differences: titrations were performed in typically in 30%, 40%, and 50% mixtures of ISA methanol:ISA water by volume to measure apparent pK$_a$ values (p$_s$K$_a$) in these mixtures. Yasuda-Shedlovsky extrapolation was subsequently used to estimate the pK$_a$ value at 0% cosolvent (Figure 5) [31, 52, 53].

$$\mathrm{psKa} + \log[\mathrm{H_2O}] = A/\epsilon + B \tag{1}$$

Yasuda-Shedlovsky extrapolation relies on the linear correlation between $\mathrm{psKa} + \log[\mathrm{H_2O}]$ and the reciprocal dielectric constant of the cosolvent mixture ($1/\epsilon$). In Eq. 1, A and B are the slope and intercept of the line fitted to experimental datapoints. Depending on the solubility requirements of the analyte, the methanol ratio of the cosolvent mixtures was adjusted. We designed the experiments to have at least 5% cosolvent ratio difference between datapoints and no more than 60% methanol content. Calculation of the Yasuda-Shedlovsky extrapolation was performed by the Sirius T3 software using at least 3 psKa values measured in different ratios of methanol:water. Addition of methanol (80%, 0.15 M KCl) was controlled by the instrument before each titration. Three consecutive pH titrations at different methanol concentrations were performed

354 using the same sample solution. In addition, three replicate measurements with independent samples
355 (prepared from the same DMSO stock) were collected.

Calculation of uncertainty in pK$_a$measurements

357 Experimental uncertainties were reported as the standard error of the mean (SEM) of three replicate
358 pK$_a$measurements. The standard error of the mean (SEM) was estimated as

$$\text{SEM} = \frac{\sigma}{\sqrt{N}} \; ; \; \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \; ; \; \mu = \frac{1}{N}\sum_{i=1}^{N}x_i \qquad (2)$$

359 where σ denotes the sample standard deviation and μ denotes the sample mean. x_i are observations and N
360 is the number of observations.

361 Since the Sirius T3 software reports pK$_a$values to only two decimal places, we have reported the SEM
362 as 0.01 in cases where SEM values calculated from 3 replicates were lower than 0.01. SEM calculated from
363 replicate measurements were found to be larger than non-linear fit error reported by the Sirius T3 Refine
364 Software from UV-absorbance model fit of a single experiment, thus leading us to believe that running
365 replicate measurements and reporting mean and SEM of pK$_a$measurements is better for capturing all
366 sources of experimental uncertainty. Notably, for UV-metric measurements, the measured pK$_a$values should
367 be insensitive to final analyte concentration and any uncertainty in the exact analyte concentration of the
368 original DMSO stock solution, justifying the use of the same stock solution (rather than independently
369 prepared stock solutions) for multiple replicates.

Quality control for chemicals

371 Compound purity was assessed by LC-MS using an Agilent HPLC 1200 Series equipped with auto-sampler,
372 UV diode array detector, and a Quadrupole MS detector 6140. ChemStation version C01.07SR2 was used
373 to analyze LC & LC/MS. An Ascentis Express C18 column (3.0 x 100 mm, 2.7 μm) was used, with column
374 temperature 45°C.

375 • Mobile phase A: 2 mM ammonium formate (pH = 3.5) aqueous
376 • Mobile phase B: 2 mM ammonium formate in acetonitrile:water = 90:10 (pH = 3.5)
377 • Flow rate : 0.75 ml/min
378 • Gradient: Starting with 10% B to 95% B in 10 minutes then hold at 95% B for 5 minutes.
379 • Post run length: 5 minutes
380 • Mass condition: ESI positive and negative mode
381 • Capillary voltage: 3000 V
382 • Drying gas flow: 12 ml/min
383 • Nebulizer pressure: 35 psi
384 • Drying temperature: 350°C
385 • Mass range: 5-1350 Da; Fragmentor: 70; Threshold: 100

386 The percent area for the primary peak is calculated based on the area of the peak divided by the total
387 area of all peaks. The percent area of the primary peak is reported as an estimate of sample purity. The
388 purity of primary LC peak was checked with ChemStation software with threshold 995, to check that there is
389 no significant impurity underneath the main peak.

NMR determination of protonation microstates

391 In general, the chemical shifts of nuclear species observed in Nuclear Magnetic Resonance (NMR) spectra
392 report on and are very sensitive to the chemical environment. Consequently, small changes in chemical
393 environment, such as the protonation events described in this work, are manifest as changes in the chemical
394 shift(s) of the nuclei. If perturbation occurs at a rate, which is fast on the NMR timescale, an average chemical
395 shift is observed. This phenomena has been exploited and utilized as a probe for determining the order
396 of protonation for molecules with more than one titratable site [54]. In some cases, direct observation of
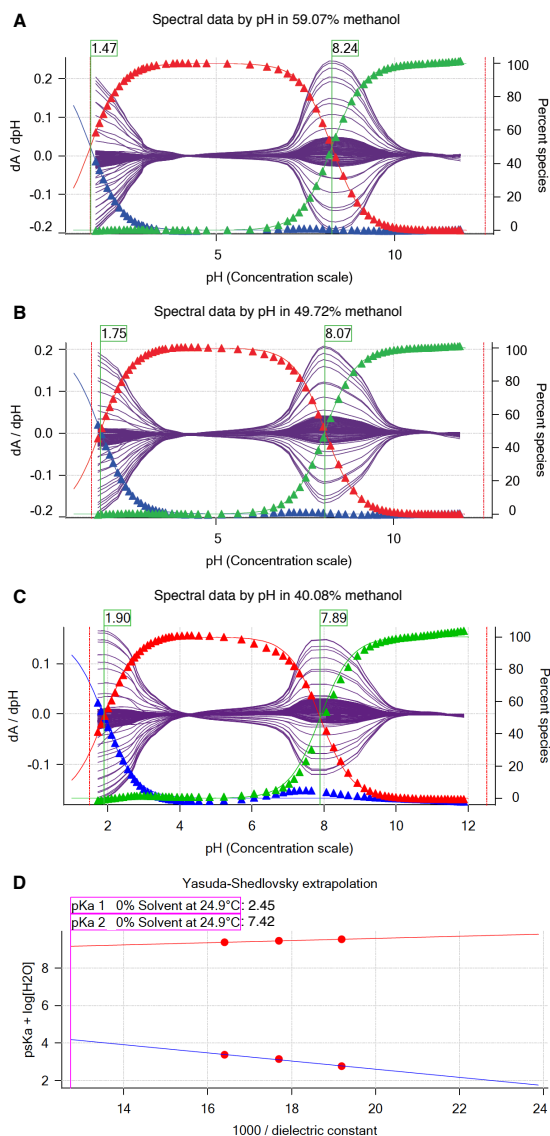397 the titrated nuclei can be difficult, for example nitrogen and oxygen, due to sample limitations and/or low

Figure 5. Determination of SM22 pK$_a$ values with cosolvent method and Yasuda-Shedlovsky extrapolation. A, **B**, and **C** show psKa of SM22 determined at various methanol concentrations: 59.07%, 49.72%, 40.08% by weight. Purple solid lines indicate derivative of absorbance with respect to pH vs pH in multiple wavelength. psKa values (green flags) were determined by Sirius T3 Refine Software. Blue, red and green triangles show relative populations of macroscopic protonation states with respect to pH calculated from the experimental data. Notice that as cosolvent concentration increases, psKa1 value decreases from 1.90 to 1.47 and psKa2 value increases from 7.84 to 8.24. **D** Yasuda-Shedlovsky extrapolation plot for SM22. Red datapoints correspond to psKa determined at various cosolvent ratios. Based on linear fitting to $psKa + log[\mathrm{H_2O}]$ vs $1/\epsilon$, pK$_{a1}$ and pK$_{a2}$ in 0% cosolvent (aqueous solution) was determined as 2.45 and 7.42, respectively. R^2 values of linear fits are both 0.99. The slope of Yasuda-Shedlovsky extrapolation shows if the observed titration has acidic (positive slope) or basic (negative slope) character dominantly, although this is an macroscopic observation and should not be relied on for annotation of pK$_a$s to functional groups (microscopic pK$_a$s).

natural abundance of the NMR active nuclei (0.37% for ^{15}N and 0.038% for ^{17}O) – amongst other factors. In these situations, chemical shifts changes of the so-called "reporter" NMR nuclei – ^{1}H, ^{31}P or ^{13}C nuclei, which are directly attached to or are a few bonds away from the titrated nuclei – have been utilized as the probe for NMR-pH titrations [21, 55, 56]. This approach is advantageous since the sensitive NMR nuclides (^{1}H and ^{31}P) are observed. In addition, ^{31}P and ^{13}C offers large spectral widths of ~300 ppm and ~200 ppm, respectively, which minimizes peak overlap. However, reporter nuclei chemical shifts provide indirect information subject to interpretation. In complex systems with multiple titratable groups, such analysis will be complicated due to a cumulative effect of these groups on the reporter nuclide due to their close proximity or the resonance observed in aromatic systems. In contrast, direct observation of the titratable nuclide where possible, affords a more straight-forward approach to studying the protonation events. In this study, the chemical shifts of the titratable nitrogen nuclei were observed using the ^{1}H-^{15}N-HMBC (Heteronuclear Multiple-Bond Correlation) experiments – a method that affords the observation of ^{15}N chemical shifts while leveraging the sensitivity accrued from the high abundance the ^{1}H nuclide.

The structures of samples SM07 and SM14 were assigned via a suite of NMR experiments, which included ^{1}H NMR, ^{13}C NMR, homonuclear correlated spectroscopy (^{1}H-^{1}H COSY), heteronuclear single quantum coherence (^{1}H-^{13}C HSQC), ^{13}C heteronuclear multiple-bond correlation (^{1}H-^{13}C-HMBC) and ^{15}N heteronuclear multiple-bond correlation (^{1}H-^{15}N-HMBC) – see SI. All NMR data used in this analysis were acquired on a Bruker 500 MHz spectrometer equipped with a 5 mm TCI CryoProbe^{TM} Prodigy at 298 K. The poor solubility of the analytes prevented analysis in water and thus water-d_2/methanol-d_4 mixture and acetonitrile-d_3 were used as solvents. The basic sites were then determined by titration of the appropriate solutions of the samples with equivalent amounts of deutero-trifluoroacetic acid (TFA-d) solution.

SM07

5.8 mg of SM07 was dissolved in 600 μL of methanol-d_4:water-d_2 (2:1 v/v ratio). A 9% v/v TFA-d solution in water-d2 was prepared, such that each 20 μL volume contained approximately 1 equivalent of TFA-d with respect to the base. SM07 solution was then titrated with the TFA-d solution at 0.5, 1.0, 1.5 and 5.0 equivalents with a ^{1}H-^{15}N HMBC spectra (optimized for 5 Hz) acquired after each TFA addition. A reference ^{1}H-^{15}N HMBC experiment was first acquired on the SM07 solution prior to commencement of the titration.

SM14

5.5 mg of SM14 was dissolved in 600 μL of acetonitrile-d_3. A 10% v/v TFA-d solution in acetonitrile-d_3 was prepared, 20 μL of which corresponds to 1 equivalent of TFA-d with respect to the base. Further 1:10 dilution of the TFA-d solution in acetonitrile-d_3, allowed measurement of 0.1 equivalent of TFA-d per 20 μL of solution. SM14 solution was then titrated with the TFA-d solutions at 0.0, 0.5, 1.0, 1.1, 1.2, 1.3, 1.5, 1.8, 2.0, 2.1, 2.6, 5.1 and 10.1 equivalents. The chemical shift changes were monitored by the acquisition of ^{1}H-^{15}N HMBC spectra (optimized for 5 Hz) after each TFA addition.

Results

Spectrophotometric pK$_a$ measurements

Spectrophotometrically-determined pK$_a$ values for all molecules from the SAMPL6 pK$_a$ challenge are shown in Figure 6 and Table 1. The protocol used—cosolvent or aqueous UV-metric titration—is indicated in Table 1 together with SEM of each reported measurement. Out of 24 molecules successfully assayed, five molecules have two resolvable pK$_a$ values, while one has three resolvable pK$_a$ values within the measurable range 2–12 pK$_a$ range. The SEM of reported pK$_a$ measurements is low, with the largest uncertainty reported being 0.04 pK units (pK$_{a1}$ of SM06 and pK$_{a3}$ of SM18). Individual replicate measurements can be found in Supplementary Table 3. Reports generated for each pK$_a$ measurement by the Sirius T3 Refine software can also be found in the Supplementary Information. Experimental pK$_a$ values for nearly all compounds with multiple resolvable pK$_a$s are well-separated (more than 3 pK$_a$ units), except for SM14 and SM18.
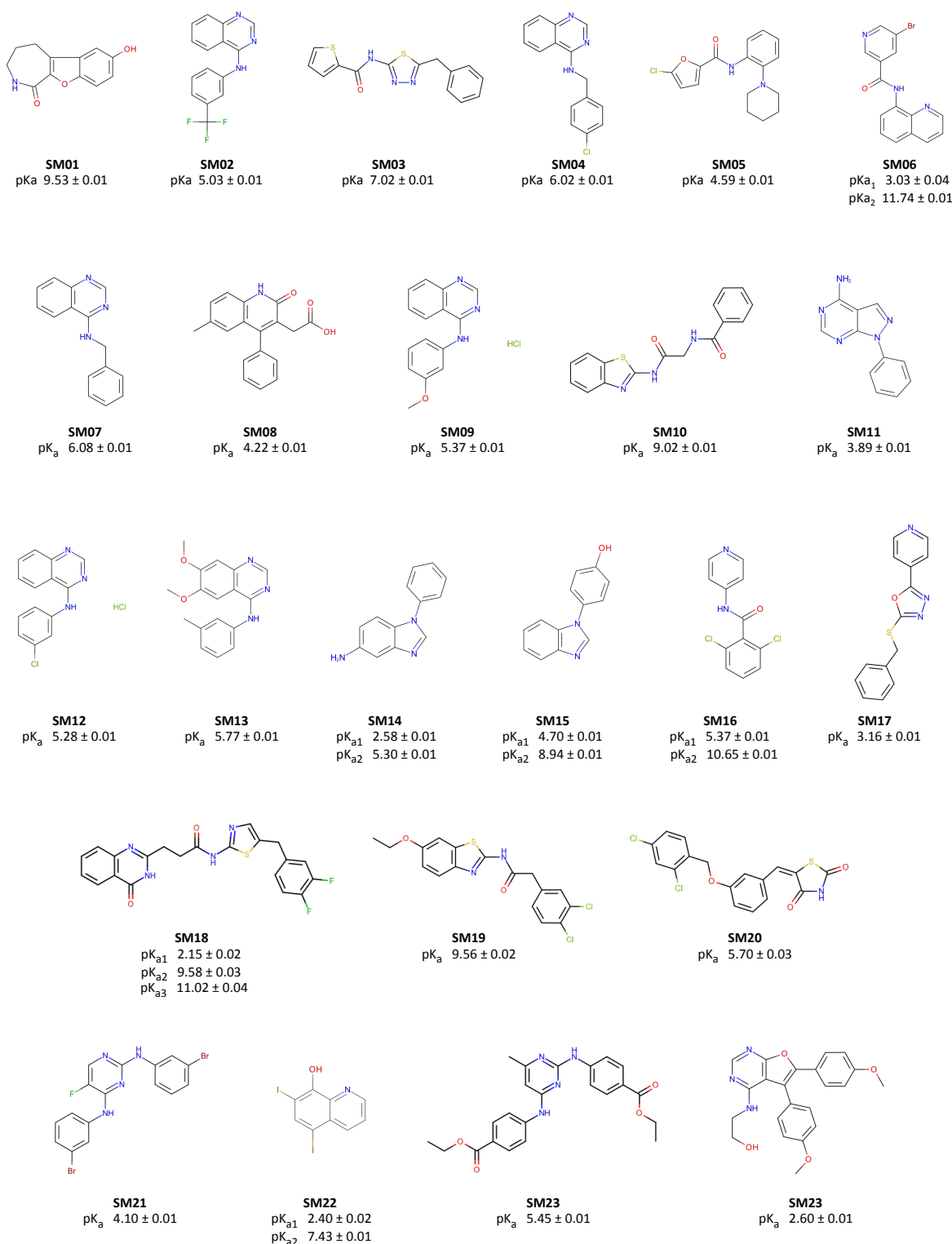
SM01
pKa 9.53 ± 0.01

SM02
pKa 5.03 ± 0.01

SM03
pKa 7.02 ± 0.01

SM04
pKa 6.02 ± 0.01

SM05
pKa 4.59 ± 0.01

SM06
pKa$_1$ 3.03 ± 0.04
pKa$_2$ 11.74 ± 0.01

SM07
pK$_a$ 6.08 ± 0.01

SM08
pK$_a$ 4.22 ± 0.01

SM09
pK$_a$ 5.37 ± 0.01

SM10
pK$_a$ 9.02 ± 0.01

SM11
pK$_a$ 3.89 ± 0.01

SM12
pK$_a$ 5.28 ± 0.01

SM13
pK$_a$ 5.77 ± 0.01

SM14
pK$_{a1}$ 2.58 ± 0.01
pK$_{a2}$ 5.30 ± 0.01

SM15
pK$_{a1}$ 4.70 ± 0.01
pK$_{a2}$ 8.94 ± 0.01

SM16
pK$_{a1}$ 5.37 ± 0.01
pK$_{a2}$ 10.65 ± 0.01

SM17
pK$_a$ 3.16 ± 0.01

SM18
pK$_{a1}$ 2.15 ± 0.02
pK$_{a2}$ 9.58 ± 0.03
pK$_{a3}$ 11.02 ± 0.04

SM19
pK$_a$ 9.56 ± 0.02

SM20
pK$_a$ 5.70 ± 0.03

SM21
pK$_a$ 4.10 ± 0.01

SM22
pK$_{a1}$ 2.40 ± 0.02
pK$_{a2}$ 7.43 ± 0.01

SM23
pK$_a$ 5.45 ± 0.01

SM23
pK$_a$ 2.60 ± 0.01

Figure 6. Molecules used in the SAMPL6 pK$_a$ challenge. Experimental UV-metric pK$_a$ measurements were performed for these 24 molecules and discernable macroscopic pK$_a$s are reported. Uncertainties are expressed as the standard error of the mean (SEM) of three independent measurements. We depicted neutral states of the molecules as sites of protonation were not determined by UV-metric methods. 2D structures were created with OpenEye OEDepict Toolkit [57]

Table 1. Experimental pK$_a$s of SAMPL6 compounds. Spectrophotometric pK$_a$ measurements were performed with two assay types based on aqueous solubility of analytes. "UV-metric pK$_a$" assay indicates spectrophotometric pK$_a$ measurements done with Sirius T3 in ISA water. "UV-metric pK$_a$ with cosolvent" assay refers to pK$_a$ determination by Yasuda-Shedlovsky extrapolation from psKa measurements in various ratios of ISA methanol:water mixtures. Triplicate measurements were performed at 25 ± 0.5 °C and in the presence of approximately 150 mM KCl to adjust ionic strength.

Molecule ID	pK$_{a1}$	pK$_{a2}$	pK$_{a3}$	Assay Type
SM01	9.53 ± 0.01			UV-metric pK$_a$
SM02	5.03 ± 0.01			UV-metric pK$_a$ with cosolvent
SM03	7.02 ± 0.01			UV-metric pK$_a$ with cosolvent
SM04	6.02 ± 0.01			UV-metric pK$_a$
SM05	4.59 ± 0.01			UV-metric pK$_a$ with cosolvent
SM06	3.03 ± 0.04	11.74 ± 0.01		UV-metric pKa
SM07	6.08 ± 0.01			UV-metric pK$_a$
SM08	4.22 ± 0.01			UV-metric pK$_a$
SM09	5.37 ± 0.01			UV-metric pK$_a$ with cosolvent
SM10	9.02 ± 0.01			UV-metric pK$_a$ with cosolvent
SM11	3.89 ± 0.01			UV-metric pK$_a$
SM12	5.28 ± 0.01			UV-metric pK$_a$
SM13	5.77 ± 0.01			UV-metric pK$_a$
SM14	2.58 ± 0.01	5.30 ± 0.01		UV-metric pK$_a$
SM15	4.70 ± 0.01	8.94 ± 0.01		UV-metric pK$_a$
SM16	5.37 ± 0.01	10.65 ± 0.01		UV-metric pK$_a$
SM17	3.16 ± 0.01			UV-metric pK$_a$
SM18	2.15 ± 0.02	9.58 ± 0.03	11.02 ± 0.04	UV-metric pK$_a$ with cosolvent
SM19	9.56 ± 0.02			UV-metric pK$_a$ with cosolvent
SM20	5.70 ± 0.03			UV-metric pK$_a$ with cosolvent
SM21	4.10 ± 0.01			UV-metric pK$_a$ with cosolvent
SM22	2.40 ± 0.02	7.43 ± 0.01		UV-metric pK$_a$ with cosolvent
SM23	5.45 ± 0.01			UV-metric pK$_a$ with cosolvent
SM24	2.60 ± 0.01			UV-metric pK$_a$ with cosolvent

[1] pK$_a$ values are reported as mean $\pm$ SEM of three replicates.

Impact of cosolvent to UV-metric pK$_a$ measurements

For molecules with insufficient aqueous solubilities throughout the titration range (pH 2-12), we resorted to cosolvent UV-metric pK$_a$ measurements, with methanol used as cosolvent. To confirm that cosolvent UV-metric pK$_a$ measurements led to indistinguishable results compared to aqueous UV-metric measurements, we collected pK$_a$ values of 12 highly soluble SAMPL6 compounds—as well as pyridoxine—using both cosolvent and aqueous methods. Correlation analysis of pK$_a$ values determined with both methods demonstrated that using methanol as cosolvent and determining aqueous pK$_a$s via Yasuda-Shedlovsky extrapolation did not result in significant bias (Figure 7), since 95% CI for mean deviation (MD) between two measurements includes 0.00 value. Means and standard errors of UV-metric pK$_a$ measurements with and without cosolvent are provided in Supplementary Table 5. pK$_a$ measurement results of individual replicate measurements with and without cosolvent can be found in Supplementary Table 4.

Purity of SAMPL6 compounds

LC-MS based purity measurements showed that powder stocks of 23 of the SAMPL6 pK$_a$ challenge compounds were >90% pure, while purity of SM22 was 87%—the lowest in the set (Supplementary Table 6). Additionally, molecular weights detected by LC-MS method were consistent with those reported in eMolecules, as well
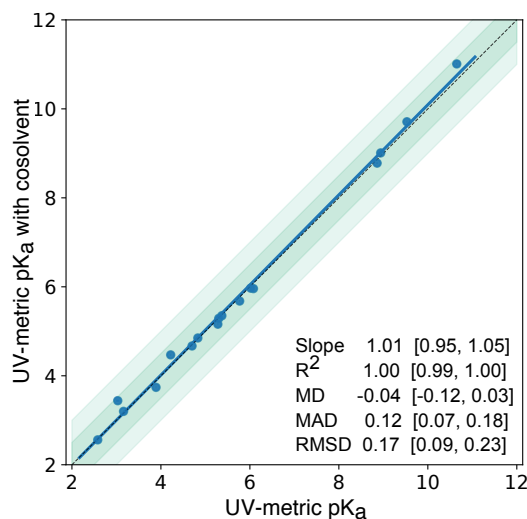
Figure 7. pK_a measurements with UV-metric method with cosolvent and UV-metric method in water show good correlation. 17 pK_a values (blue marks) of 13 chemicals were measured with both UV-metric pK_a method in water and UV-metric pK_a method with methanol as cosolvent (Yasuda-Shedlovsky extrapolation to 0% methanol). Dashed black line has slope of 1, representing perfect correlation. Dark and light green shaded areas indicate ± 0.5 and ± 1.0 pK_a unit difference regions, respectively. Error bars are plotted as SEM of replicate measurements, although they are not visible since the largest SEM value is 0.04. MD: Mean difference, MAD: Mean absolute deviation, RMSD: Root-mean-square deviation. Statistic are reported with 95th percentile confidence intervals calculated over 10000 bootstrap samples.

458 as supplier-reported molecular weights, when provided. It is recommended to use compounds with ~90%
459 purity to minimize the impact on high-accuracy pK_a measurements. Impurities with no UV-chromophore,
460 or elute too late in LC may not be detected with this method, although chances are small. The peak purity
461 check of primary peak can detect the presence of a large impurity underneath the main peak, but if the UV
462 spectrum of the impurity is exactly same with analyte in the main peak, it may not be resolved. HPLC UV
463 detector's wavelength inaccuracy is <1%. Mass inaccuracy of MS instrument is 0.13 u within the calibrated
464 mass range in the scan mode.

Characterization of SM07 microstates with NMR

466 ^{15}N Chemical shifts (ppm, referenced to ammonia at 0 ppm) for N-8, N-10 and N-12 – measured from
467 the ^{1}H-^{15}N HMBC experiments – were plotted against the titrated TFA-*d* equivalents (0.0, 0.5, 1.0, 1.5,
468 and 5.0 equivalents)(Figure 8 A). A large upfield shift of ~82 ppm is observed for N-12. The initial linear
469 relationship between chemical shift and TFA equivalents, shown in Figure SI 25 for N-12, is expected for
470 strong monoprotic bases – as is the case for SM07. The large upfield chemical shift change (82 ppm) is
471 consistent with a charge delocalization as shown in the resonance structures inset of Figure SI 25. Further
472 evidence for this delocalization is observed for N-8, which shows a downfield chemical shift change of ~28
473 ppm compared to just ~1.5 ppm for N-10. Titration of SM07 with more than 1 equivalents of TFA-*d* did not
474 result in further significant chemical shift changes – establishing that SM07 is a monoprotic base(Figure 8A).

Characterization of SM14 microstates with NMR

476 Determining the protonation sites for SM14, which has pK_a values of 2.58 and 5.30 (Table 1), was more
477 challenging due to multiple possible resonance structures in the mono- and di-protonated states. We noticed
478 that water/methanol co-solvent exhibited strong solvent effects, which complicated the data interpretation
479 for SM14. For instance, titration of SM14 in methanol/water (see SI) showed incomplete protonation of
480 N-9 even after 5 equivalents of TFA-d were added. This observation is consistent with UV-metric psKa

measurements done in methanol cosolvent, where both psKa values were decreasing as the percentage of methanol was increased, making observation of these protonation states more difficult. Thus the utilization of an aprotic solvent was necessary for unambiguous interpretation of the data.

Acetonitrile-d_3 was selected as our solvent of choice. Titration of SM14 (5.5 mg) with up to 10 equivalents of TFA-d in acetonitrile-d_3 (0.0, 0.5, 1.0, 1.1, 1.2, 1.3, 1.5, 1.8, 2.0, 2.1, 2.6, 5.1, and 10.1 equivalents), provided a much clearer picture of its protonation states((Figure 8 B)). N-9, with the large upfield chemical shift change ~72 ppm at 1 equivalent of TFA-d, clearly is the site of first protonation. At this point, the downfield chemical shift changes observed for N-7 ($\Delta\delta \approx 6.5$) and N-16 ($\Delta\delta \approx 5$) can be attributed to electronic effects rather than a direct protonation. The large upfield shift for N-9 indicates this to be the site of first protonation; complete protonation was attained at roughly 2.5 equivalents of TFA-d, suggesting that SM14 is a weak base under these experimental conditions. Following the protonation of N-9, a second protonation event occurs at N-16 nitrogen as evident by the upfield chemical shift change observed for N-16. However, a continuous change in the chemical shift of N-16 even after addition of 10 equivalents of TFA-d indicates that this protonation event is incomplete but provides evidence for N-16 being the second protonation site. This observation is consistent with N-16 being even a weaker base than N-9, which is expected of the aniline-type amines. Other notable observations are the slight downfield chemical shift changes for N-7 and N-9, during the second protonation event. This is attributed to electronic effects from the protonation of N-16.

Discussion

Sample preparation and effect of cosolvents in UV-metric measurements

Samples for UV-metric pK$_a$ measurements were prepared by dilution of up to 5 µL DMSO stock solution of analyte in 1.5 mL ISA water, which results in the presence of ~0.3% DMSO during titration, which has negligible effect on pK$_a$ measurements. For UV-metric or pH-metric measurements, it is possible to prepare samples without DMSO, but it is difficult to prepare samples by weighing extremely low amounts of solid stocks (in the order of 0.01–0.10 mg) to target 50 µM analyte concentrations, even with an analytical balance. For experimental throughput, we therefore preferred using DMSO stock solutions. Another advantage of starting from DMSO stock solutions is that it helps to overcome kinetic solubility problems of analytes.

In UV-metric measurements, both water and methanol (when used as cosolvent) stock solutions were ionic strength adjusted with 150 mM KCl, but acid and base solutions were not. This means that throughout pH titration ionic strength slighly fluctuates, but on average ionic strength of samples were staying around 150–180 mM. By using ISA solutions the effect of salt concentration change on pK$_a$ measurements was minimized.

If an analyte is soluble enough, UV-metric pK$_a$ measurements in water should be preferred over co-solvent methods, since pK$_a$ measurement in water is more direct. For pK$_a$ determination via cosolvent extrapolation using methanol, apparent pK$_a$s (p$_s$K$_a$) in at least three different methanol:water ratios must be measured, and pK$_a$ in 0% cosolvent computed by Yasuda-Shedlovsky extrapolation. The number and spread of p$_s$K$_a$ measurements and error in linear fit extrapolation influences the accuracy of pK$_a$s determined by this approach. To test that UV-metric methods with or without cosolvent have indistinguishable performance, we collected pK$_a$ values for 17 SAMPL6 compounds and pyridoxine with both methods. Figure 7 shows there is good correlation between both methods and the mean absolute deviation between two methods is 0.12 (95% CI [0.07, 0.18]). Mean deviation between the two sets is -0.04 (95% CI [-0.12, 0.03]), showing there is no bias in cosolvent measurements. The largest absolute deviation observed was 0.41 for SM06.

Impact of impurities to UV-metric pK$_a$ measurements

Precisely how much the presence of small amounts of impurities impact UV-metric pK$_a$ measurements is unpredictable. For an impurity to alter UV-metric pK$_a$ measurements, it must possess a UV-chromophore and a titratable group in the vicinity of the chromophore—otherwise, it would not interfere with absorbance signal of the analyte. If a titratable impurity *does* possess a UV-chromophore, UV multiwavelength absorbance from the analyte and impurity will be convoluted. The impact the impurity will have on the multiwavelength absorbance spectra depends on the strength of its molar absorption coefficient. In the worst case scenario,
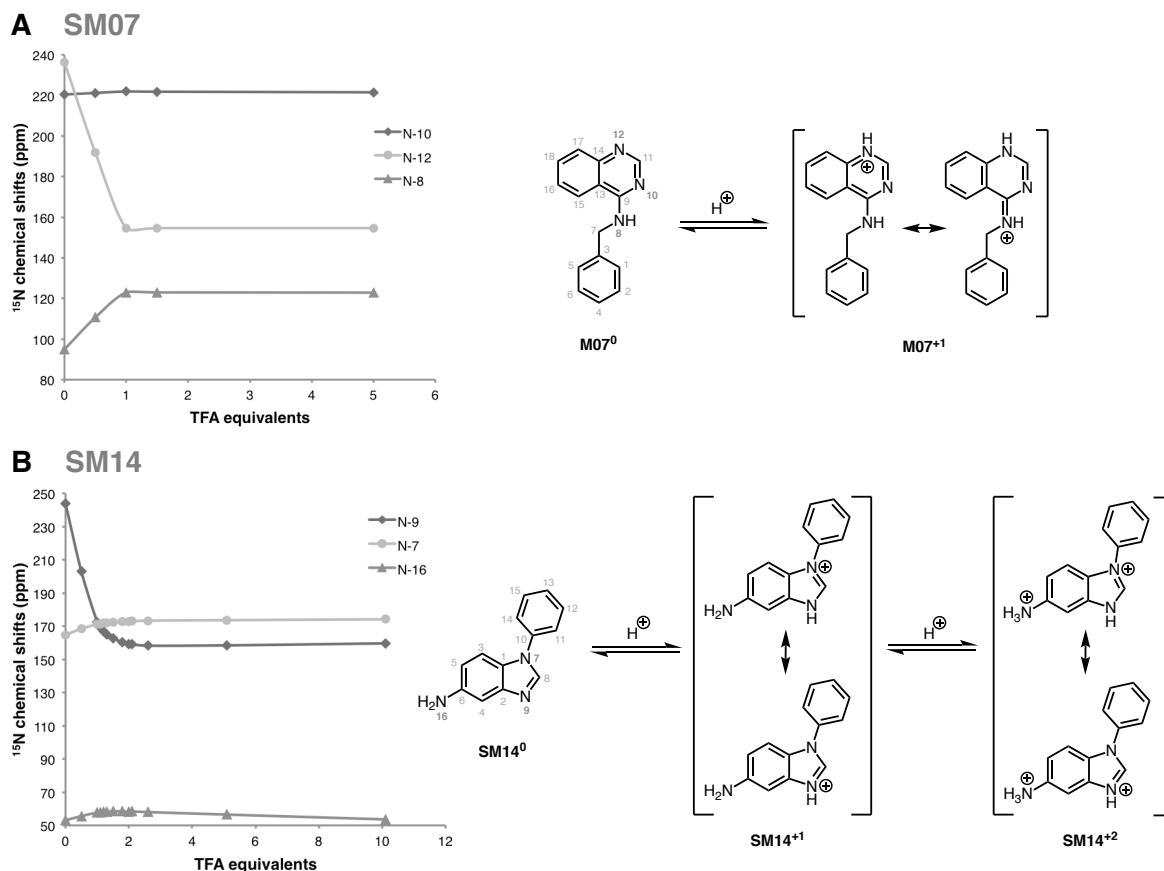
Figure 8. Dominant protonation microstates of SM07 and SM14 characterized by NMR. (A) Sequence of protonation sites of SM07 were determined by ^{1}H-^{15}N HMBC experiments in 1:2 water:methanol mixture. *Left:* The plot of ^{15}N chemical shifts of N-10, N-12, and N-8 of SM07 *vs* titrated TFA-*d* equivalents, showing the mono-protonation of N-12 as evidenced by its large upfield chemical shifts change. Acidity of the medium increases as more equivalents of TFA-*d* were added. Electronic effects due to protonation of N-12 cause downfield chemical shift change of N-10 and N-8 between 0-1 equivalents of TFA-*d*. *Right:* NMR based model of the order of dominant protonation states for M07. Only one protonation was observed at N-12. Microstates shown in the figure are the most likely contributors to the UV-metric pK$_a$value 6.08 ± 0.01. **(B)** Sequence of protonation sites of SM14 were determined by ^{1}H-^{15}N HMBC experiments in acetonitrile. *Left:* The plot of ^{15}N chemical shifts of N-9, N-7, and N-16 of SM14 *vs* titrations of TFA-*d* equivalents, showing two sequential protonation events. The first protonation at N-9 with a large upfield chemical shift change of 71.6 ppm was seen between 0 and 1 equivalents of TFA-*d*. Downfield chemical shift change of N-7 and N-19 in this region were due the electronic effect from the protonation of N-9. N-16 also exhibited a small upfield chemical shift change of 4.4 ppm between 2.5 and 10 equivalents of TFA-*d*, which indicated N-16 as the second site of protonation. *Right:* NMR based model of the order of dominant protonation states for M14, showing two sequential protonation events. Also, two pK$_a$values were detected with UV-metric pK$_a$measurements for SM14. Assuming sequence of protonation events will be conserved between water and acetonitrile solvents, SM14[0] and SM14[+1] microstates shown in the figure are the major contributors to the UV-metric pK$_a$value 5.30 ± 0.01. SM14[+1] and SM14[+2] microstates shown in the figure are the major pair of microstates contributing to the UV-metric pK$_a$value 2.58 ± 0.01. There could be minor microstates with very low populations that could not be distinguished in these NMR experiments.

529 an impurity with high concentration or strong UV absorbance can shift the measured pK_a value or create
530 the appearance of an extra pK_a. As a result, it is important to use analytes with high purities to obtain high
531 accuracy pK_a measurements. Therefore, we confirmed the purities of SAMPL6 compounds with LC-MS.

Interpretation of UV-metric pK_a measurements

533 Multiwavelength absorbance analysis on the Sirius T3 allows for very good resolution of pK_as, but it is
534 important to note that this method produces *macroscopic* pK_as for polyprotic compounds. If multiple
535 microscopic pK_as have close pK_a values and overlapping changes in UV absorbance spectra associated with
536 protonation/deprotonation, the spectral analysis could produce a single macroscopic pK_a that represents
537 an aggregation of multiple microscopic pK_as. An extreme example of such case is demonstrated in the
538 simulated macrostate populations of cetirizine that would be observed with UV-metric titration (Figure 2).

539   If protonation state populations observed via UV-metric titrations (such as in Figure 3B) are composed
540 of a single microstate, experimentally measured pK_as are indeed microscopic pK_as. Unfortunately, judging
541 the composition of experimental populations is not possible by just using UV-metric or pH-metric titration.
542 Molecules in the SAMPL6 pK_a challenge dataset with only one pK_a value measured in the 2–12 range could
543 therefore be monoprotic (possessing a single titratable group that changes protonation state by gain or
544 loss of a single proton over this pH range) or polyprotic (gaining or losing multiple protons from one or
545 more sites with overlapping microscopic pK_a values). Similarly, titration curves of molecules with multiple
546 experimental pK_as may show well-separated microscopic pK_as or macroscopic experimental pK_as that are
547 really composites of microscopic pK_as with similar values. Therefore, without additional experimental
548 evidence, UV-metric pK_as should not be assigned to individual titratable groups.

549   Determination of the exact microstates populated at different pH values via NMR can provide a com-
550 plementary means of differentiating between microscopic and macroscopic pK_as in cases where there is
551 ambiguity. As determination of protonation microstates via NMR is very laborious, we were only able to
552 characterize microstates of two molecules: SM07 and SM11.

553   In UV-metric pK_a measurements with cosolvent, the slope of Yasuda-Shedlovsky extrapolation can be
554 interpreted to understand if the pK_a has dominantly acidic or basic character. As the methanol ratio is
555 increased, psKa values of acids increase, while psKa values for bases decrease. However, it is important to
556 remember that if the measured pK_a is macroscopic, acid/base assignment from cosolvent p_sK_a trends is also
557 a macroscopic property, and should not be used as a guide for assigning pK_a values to functional groups [58].

NMR microstate characterization

559 The goal of NMR characterization was to collect information on microscopic states related to experimental
560 pK_a measurements, i.e., determine exact sites of protonation. pK_a measurements performed with spec-
561 trophotometric method provide macroscopic pK_a values, but do not provide information on the site(s) of
562 protonation. On the other hand, most computational prediction methods primarily predict microscopic
563 pK_a values. Protonation sites can be determined by NMR methods, although these measurements are very
564 laborious in terms of data collection and interpretation compared to pK_a measurements with the automated
565 Sirius T3. Moreover, not all SAMPL6 molecules were suitable for NMR measurements due to the high
566 sample concentration requirements (for methods other than proton NMR, such as ^{13}C and ^{15}N spectra) and
567 limiting analyte solubility. Heavy atom spectra that rely on natural isotope abundance require high sample
568 concentrations (preferably in the order of 100 mM). It is possible that drug or drug-fragment-like compounds,
569 such as the compounds used in this study, have insufficient aqueous solubility, limiting the choice of solvent
570 and pH. It may be necessary to use organic cosolvents to prepare these high concentration solutions, or
571 only prepare samples at pH values that correspond to high solubility states (e.g., when charged state of the
572 compound is populated).

573   We performed NMR based microstate characterization only for SM07 and SM14. We were able to identify
574 the order of dominant protonation microstates, as shown in Figure 8. These pairs of microstates and
575 the order of microscopic transitions can be associated with experimental pK_as determined by UV-metric
576 titrations, under the assumption that different organic solvents used in NMR measurements will have
577 negligible effect on the sequence of microstates observed as the medium was titrated with acid, although

578 shift in pK$_a$ values is expected. NMR measurements for SM07 and SM14 were done in water:methanol(1:2
579 v/v)) and acetonitrile solutions, respectively. On the other hand, pK$_a$ values of these two compounds
580 were determined by UV-metric titrations in ISA water. Additional UV-metric pK$_a$ measurements of these
581 compounds with methanol as cosolvent showed that their psKa values decreased as the cosolvent ratio
582 increased (i.e. dielectric constant decreased) as expected from base titration sites. Identification of SM07
583 and SM14 titratable sites type as base is consistent between NMR based models and UV-metric cosolvent
584 titrations. The order of microstates observed in the titration of NMR samples are very likely to corresponds
585 to the dominant microstates associated with UV-metric pK$_a$ measurements. N-12 of SM07 was observed
586 as the only protonation site of SM07 during TFA-d titration up to 5 equivalents which supports that SM07
587 is a mono-protic acid and UV-metric pK$_a$ value 6.08 ± 0.01 corresponds to microscopic protonation of N-12.
588 For SM14, two protonation sites were observed (N-16 and N-9, in the order if increasing psKa). Microstate
589 pairs shown in Figure 8B were determined as dominant contributors to UV-metric pK$_a$s 2.58 ± 0.01 and
590 5.30 ± 0.01, although minor microspecies with very low populations (undetected in NMR experiments) could
591 be contributing to macroscopic pK$_a$ values observed by UV-metric method.

592 In addition to SM07, there were five other 4-amino quinazoline derivatives in SAMPL6 set: SM02, SM04.
593 SM09, SM12, and SM13. For these series, all the potential titratable sites are located in 4-amino quinazoline
594 scaffold and there are no other additional titratable sites present in these compounds compared to SM07.
595 Therefore, based on structural similarity, it is reasonable to predict that N-12 is the microscopic protonation
596 site for all of these compounds. We can infer that UV-metric pK$_a$ values measured for 4-amino quinazoline
597 series are also microscopic pK$_a$s and they are related to the protonation of the same quinazoline nitrogen
598 with the same neutral background protonation states as shown for SM07 in Figure 8A.

Suggestions for future pK$_a$ data collection for computational challenges

600 Most high-throughput pK$_a$ measurement methods measure macroscopic pK$_a$s. One way to circumvent
601 this problem is to confine our interest in future pK$_a$ challenges to experimental datasets containing only
602 monoprotic compounds if UV-metric or pH-metric pK$_a$ measurements are the method of choice. Then it
603 would be possible to unambiguously assign pK$_a$ values to underlying protonation states. However, it is
604 important to consider that multiprotic compounds are common in pharmaceutically interesting molecules
605 and it is important to be able to model them reliably.

606 Although relatively efficient, UV-metric pK$_a$ measurements with the Sirius T3 do not provide structural
607 information about microstates. Even the acid-base assignment based on direction of psKa shift with
608 cosolvent is not a reliable indicator for assigning experimental pK$_a$ values to individual functional groups
609 in in multiprotic compounds. On the other hand, most computational pK$_a$ prediction methods output
610 microscopic pK$_a$s. It is therefore difficult to use experimental macroscopic pK$_a$ values to assess and train
611 microscopic pK$_a$ prediction methods directly without further means of annotating macroscopic-microscopic
612 correspondence. It is usually impossible to infer the underlying microscopic pK$_a$ values from macroscopic
613 measurements of a polyprotic compound without complementary experiments that can provide structural
614 information. Therefore, for future data collection efforts for evaluation of pK$_a$ predictions, if measurement of
615 pK$_a$s via NMR is not possible, we advise supplementing UV-metric measurements with NMR characterization
616 of microstates to show if observed pK$_a$s are microscopic (related to a single group) or macroscopic (related
617 to dissociation of multiple groups), as performed for SM07 and SM14 in this study.

618 Another source of complexity is how the observed macroscopic pK$_a$s, i.e., the situations under which
619 microstates become indistinguishable, changes between different experimental methods as illustrated
620 in Figure 2. The "macroscopic" label is commonly ascribed to transitions between different ionization
621 states of a molecule (all microstates that have the same total charge form one macrostate), but this
622 definition only applies to potentiometric methods. In UV-absorbance based methods, the principle that
623 determines which microstates will be distinguishable is not charge or number of bound protons, but
624 molecular absorbance changes, and how closely underlying microscopic pK$_a$ values overlap. To compare
625 experimental macroscopic pK$_a$ and microscopic computational predictions on common ground, the best
626 solution is to compute "predicted" macroscopic pK$_a$ values from microscopic pK$_a$s following the principle of
627 experiment type. A disadvantage of this approach is that experimental data cannot provide direct guidance

on microscopic pK_a resolution for improving pK_a prediction methods.

Since analyte purity is critical for accuracy, necessary quality control experiments must be performed to ensure at least 90% purity for UV-metric pK_a measurements. Higher purity values may be necessary for other methods. For potentiometric methods, knowing the stoichiometry of any counterions present in the original powder stocks is also necessary.

For the set of SAMPL6 pK_a challenge compounds, we could not use potentiometric pK_a measurements due to the low aqueous solubility of many of these compounds. The lowest solubility observed *somewhere* in the experimental pH range of titration is the limiting factor, since for accurate measurements the analyte must stay in the solution phase throughout the entire titration. Since the titration pH range is determined with the goal of capturing all ionization states, the analyte is inevitably exposed to pH values that correspond to low solubility. Neutral and zwitterionic species can be orders of magnitude less soluble than ionic species.

For future pK_a challenges with multiprotic compounds, if sufficient time and effort can be spared, it would be ideal to construct an experimental pK_a dataset using experimental methods that can measure microscopic pK_as directly, such as NMR. In this study, we were only able to perform follow up NMR microstate characterization of two compounds due to the labor-intensive nature of NMR data collection and analysis.

In the future pK_a challenges, it would be especially interesting to expand this exercise to larger and more flexible drug-like molecules. pK_a values are environment dependent and it would be useful to be able to predict pK_a shifts based on on ionic strength, temperature, lipophilic content, with cosolvents or in organic solvents. Measuring the pK_a of molecules in organic solvents would be useful for guiding process chemistry. To test such predictions, special pK_a experiments would need to be designed to measure pK_as under different conditions.

Code and data availability

- SAMPL6 pK_a challenge instructions, submissions, experimental data and analysis is available at https://github.com/MobleyLab/SAMPL6
- Python scripts used for compound selection are available at **compound_selection** directory of https://github.com/choderalab/sampl6-physicochemical-properties

Overview of supplementary information

Organized in SI document:

- TABLE SI 1: procurement details of SAMPL6 compounds
    - TABLE SI 2: selection details of SAMPL6 compounds
    - TABLE SI 3: pK_a results of replicate experiments csv
    - TABLE SI 4: pK_a results of water and cosolvent replicate experiments csv
    - TABLE SI 5: pK_a mean and SEM results of water and cosolvent replicate experiments
    - TABLE SI 6: Summary of LC-MS purity results
    - FIGURE SI 1 - 24: LC-MS Figures
    - FIGURE SI 25: NMR characterization of SM07 microstates
    - FIGURE SI 26: NMR characterization of SM014 microstates

Extra files:

- Sirius T3 reports: supplementary_files.zip

Author Contributions

Conceptualization, MI, JDC, TR, ASR, DLM ; Methodology, MI, DL, IEN ; Software, MI, ASR ; Formal Analysis, MI ; Investigation, MI, DL, IEN, HW, XW, MR; Resources, TR, DL; Data Curation, MI ; Writing-Original Draft, MI, JDC, IEN; Writing - Review and Editing, MI, DL, ASR, IEN, HW, XW, MR, GEM, DLM, TR, JDC; Visualization, MI, IEN ; Supervision, JDC, TR, DLM, GEM, AAM ; Project Administration, MI ; Funding Acquisition, JDC, DLM, TR, MI.

Acknowledgments

Disclosures

JDC is a member of the Scientific Advisory Board for Schrödinger, LLC. DLM is a member of the Scientific Advisory Board of OpenEye Scientific Software.

References

[1] **Mobley DL**, Chodera JD, Isaacs L, Gibb BC. Advancing predictive modeling through focused development of model systems to drive new modeling innovations. UC Irvine: Department of Pharmaceutical Sciences, UCI. 2016; https://escholarship.org/uc/item/7cf8c6cr.

[2] **Drug Design Data Resource**, SAMPL;. https://drugdesigndata.org/about/sampl.

[3] **Nicholls A**, Mobley DL, Guthrie JP, Chodera JD, Bayly CI, Cooper MD, Pande VS. Predicting Small-Molecule Solvation Free Energies: An Informal Blind Test for Computational Chemistry. J Med Chem. 2008 Feb; 51(4):769–779. doi: 10.1021/jm070549+.

[4] **Guthrie JP**. A Blind Challenge for Computational Solvation Free Energies: Introduction and Overview. J Phys Chem B. 2009 Jan; 113(14):4501–4507.

[5] **Skillman AG**, Geballe MT, Nicholls A. SAMPL2 Challenge: Prediction of Solvation Energies and Tautomer Ratios. J Comput Aided Mol Des. 2010 Apr; 24(4):257–258. doi: 10.1007/s10822-010-9358-0.

[6] **Geballe MT**, Skillman AG, Nicholls A, Guthrie JP, Taylor PJ. The SAMPL2 Blind Prediction Challenge: Introduction and Overview. J Comput Aided Mol Des. 2010 May; 24(4):259–279. doi: 10.1007/s10822-010-9350-8.

[7] **Skillman AG**. SAMPL3: Blinded Prediction of Host–guest Binding Affinities, Hydration Free Energies, and Trypsin Inhibitors. J Comput Aided Mol Des. 2012 May; 26(5):473–474. doi: 10.1007/s10822-012-9580-z.

[8] **Geballe MT**, Guthrie JP. The SAMPL3 Blind Prediction Challenge: Transfer Energy Overview. J Comput Aided Mol Des. 2012 Apr; 26(5):489–496. doi: 10.1007/s10822-012-9568-8.

[9] **Muddana HS**, Varnado CD, Bielawski CW, Urbach AR, Isaacs L, Geballe MT, Gilson MK. Blind Prediction of Host–guest Binding Affinities: A New SAMPL3 Challenge. J Comput Aided Mol Des. 2012 Feb; 26(5):475–487. doi: 10.1007/s10822-012-9554-1.

[10] **Guthrie JP**. SAMPL4, a Blind Challenge for Computational Solvation Free Energies: The Compounds Considered. J Comput Aided Mol Des. 2014 Apr; 28(3):151–168. doi: 10.1007/s10822-014-9738-y.

[11] **Mobley DL**, Wymer KL, Lim NM, Guthrie JP. Blind Prediction of Solvation Free Energies from the SAMPL4 Challenge. J Comput Aided Mol Des. 2014 Mar; 28(3):135–150. doi: 10.1007/s10822-014-9718-2.

[12] **Muddana HS**, Fenley AT, Mobley DL, Gilson MK. The SAMPL4 Host–guest Blind Prediction Challenge: An Overview. J Comput Aided Mol Des. 2014 Mar; 28(4):305–317. doi: 10.1007/s10822-014-9735-1.

[13] **Mobley DL**, Liu S, Lim NM, Wymer KL, Perryman AL, Forli S, Deng N, Su J, Branson K, Olson AJ. Blind Prediction of HIV Integrase Binding from the SAMPL4 Challenge. J Comput Aided Mol Des. 2014 Mar; 28(4):327–345. doi: 10.1007/s10822-014-9723-5.

[14] **Yin J**, Henriksen NM, Slochower DR, Shirts MR, Chiu MW, Mobley DL, Gilson MK. Overview of the SAMPL5 Host–guest Challenge: Are We Doing Better? J Comput Aided Mol Des. 2017; 31(1):1–19. doi: 10.1007/s10822-016-9974-4.

[15] **Bannan CC**, Burley KH, Chiu M, Shirts MR, Gilson MK, Mobley DL. Blind Prediction of Cyclohexane–water Distribution Coefficients from the SAMPL5 Challenge. J Comput Aided Mol Des. 2016 Sep; 30(11):1–18. doi: 10.1007/s10822-016-9954-8.

[16] **Bannan CC**, Burley KH, Chiu M, Shirts MR, Gilson MK, Mobley DL. Blind prediction of cyclohexane–water distribution coefficients from the SAMPL5 challenge. Journal of Computer-Aided Molecular Design. 2016 Nov; 30(11):927–944. http://link.springer.com/10.1007/s10822-016-9954-8, doi: 10.1007/s10822-016-9954-8.

[17] **Rustenburg AS**, Dancer J, Lin B, Feng JA, Ortwine DF, Mobley DL, Chodera JD. Measuring experimental cyclohexane-water distribution coefficients for the SAMPL5 challenge. Journal of Computer-Aided Molecular Design. 2016 Nov; 30(11):945–958. http://link.springer.com/10.1007/s10822-016-9971-7, doi: 10.1007/s10822-016-9971-7.

[18] **Pickard FC**, König G, Tofoleanu F, Lee J, Simmonett AC, Shao Y, Ponder JW, Brooks BR. Blind prediction of distribution in the SAMPL5 challenge with QM based protomer and pK a corrections. Journal of Computer-Aided Molecular Design. 2016 Nov; 30(11):1087–1100. http://link.springer.com/10.1007/s10822-016-9955-7, doi: 10.1007/s10822-016-9955-7.

[19] **Bodner GM**. Assigning the pKa's of polyprotic acids. J Chem Educ. 1986; 63(3):246.

[20] **Darvey IG**. The assignment of pKa values to functional groups in amino acids. Wiley Online Library; 1995.

[21] **Bezençon J**, Wittwer MB, Cutting B, Smieško M, Wagner B, Kansy M, Ernst B. pKa determination by 1H NMR spectroscopy – An old methodology revisited. Journal of Pharmaceutical and Biomedical Analysis. 2014 May; 93:147–155. http://linkinghub.elsevier.com/retrieve/pii/S0731708513005992, doi: 10.1016/j.jpba.2013.12.014.

[22] **Elson EL**, Edsall JT. Raman spectra and sulfhydryl ionization constants of thioglycolic acid and cysteine. Biochemistry. 1962; 1(1):1–7.

[23] **Elbagerma MA**, Edwards HGM, Azimi G, Scowen IJ. Raman spectroscopic determination of the acidity constants of salicylaldoxime in aqueous solution. Journal of Raman Spectroscopy. 2011 Mar; 42(3):505–511. http://doi.wiley.com/10.1002/jrs.2716, doi: 10.1002/jrs.2716.

[24] **Sober HA**, Company CR. Handbook of Biochemistry: Selected Data for Molecular Biology. Handbook of Biochemistry: Selected Data for Molecular Biology, Chemical Rubber Company; 1970. https://books.google.com/books?id=16QRAQAAMAAJ.

[25] **Benesch RE**, Benesch R. The Acid Strength of the -SH Group in Cysteine and Related Compounds. Journal of the American Chemical Society. 1955; 77(22):5877–5881. https://doi.org/10.1021/ja01627a030, doi: 10.1021/ja01627a030.

[26] **Rupp M**, Korner R, V Tetko I. Predicting the pKa of small molecules. Combinatorial chemistry & high throughput screening. 2011; 14(5):307–327.

[27] **Marosi A**, Kovács Z, Béni S, Kökösi J, Noszál B. Triprotic acid–base microequilibria and pharmacokinetic sequelae of cetirizine. European Journal of Pharmaceutical Sciences. 2009 Jun; 37(3-4):321–328. http://linkinghub.elsevier.com/retrieve/pii/S0928098709000773, doi: 10.1016/j.ejps.2009.03.001.

[28] **Tam KY**, Takács-Novák K. Multi-wavelength spectrophotometric determination of acid dissociation constants: a validation study. Analytica chimica acta. 2001; 434(1):157–167.

[29] **Allen RI**, Box KJ, Comer JEA, Peake C, Tam KY. Multiwavelength spectrophotometric determination of acid dissociation constants of ionizable drugs. Journal of pharmaceutical and biomedical analysis. 1998; 17(4):699–712.

[30] **Comer JEA**, Manallack D. Ionization Constants and Ionization Profiles. In: *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering* Elsevier; 2014.http://linkinghub.elsevier.com/retrieve/pii/B9780124095472112338, doi: 10.1016/B978-0-12-409547-2.11233-8.

[31] **Avdeef A**, Box KJ, Comer JEA, Gilges M, Hadley M, Hibbert C, Patterson W, Tam KY. PH-metric logP 11. pK a determination of water-insoluble drugs in organic solvent–water mixtures. Journal of pharmaceutical and biomedical analysis. 1999; 20(4):631–641.

[32] **Cabot JM**, Fuguet E, Rosés M, Smejkal P, Breadmore MC. Novel Instrument for Automated pK_a Determination by Internal Standard Capillary Electrophoresis. Analytical Chemistry. 2015 Jun; 87(12):6165–6172. http://pubs.acs.org/doi/10.1021/acs.analchem.5b00845, doi: 10.1021/acs.analchem.5b00845.

[33] **Wan H**, Holmén A, N\a ag\a ard M, Lindberg W. Rapid screening of pKa values of pharmaceuticals by pressure-assisted capillary electrophoresis combined with short-end injection. Journal of Chromatography A. 2002; 979(1-2):369–377.

[34] **Reijenga J**, van Hoof A, van Loon A, Teunissen B. Development of Methods for the Determination of pK_a Values. Analytical Chemistry Insights. 2013 Jan; 8:ACI.S12304. http://journals.sagepub.com/doi/10.4137/ACI.S12304, doi: 10.4137/ACI.S12304.

[35] **Sterling T**, Irwin JJ. ZINC 15 – Ligand Discovery for Everyone. Journal of Chemical Information and Modeling. 2015 Nov; 55(11):2324–2337. http://pubs.acs.org/doi/10.1021/acs.jcim.5b00559, doi: 10.1021/acs.jcim.5b00559.

[36] **Baell JB**, Holloway GA. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. Journal of Medicinal Chemistry. 2010 Apr; 53(7):2719–2740. http://pubs.acs.org/doi/abs/10.1021/jm901137j, doi: 10.1021/jm901137j.

[37] **Saubern S**, Guha R, Baell JB. KNIME Workflow to Assess PAINS Filters in SMARTS Format. Comparison of RDKit and Indigo Cheminformatics Libraries. Molecular Informatics. 2011 Oct; 30(10):847–850. http://doi.wiley.com/10.1002/minf.201100076, doi: 10.1002/minf.201100076.

[38] eMolecules Database Free Version;. Accessed: 2017-06-01. https://www.emolecules.com/info/products-data-downloads.html.

[39] OEChem Toolkit 2017.Feb.1;. OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.

[40] **Shelley JC**, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M. Epik: a software program for pK_a prediction and protonation state generation for drug-like molecules. Journal of Computer-Aided Molecular Design. 2007 Dec; 21(12):681–691. http://link.springer.com/10.1007/s10822-007-9133-z, doi: 10.1007/s10822-007-9133-z.

[41] Schrödinger Release 2016-4: Epik Version 3.8;. Schrödinger, LLC, New York, NY, 2016.

[42] OEMolProp Toolkit 2017.Feb.1;. OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.

[43] **Wishart DS**. DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Research. 2006 Jan; 34(90001):D668–D672. https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkj067, doi: 10.1093/nar/gkj067.

[44] **Pence HE**, Williams A. ChemSpider: An Online Chemical Information Resource. Journal of Chemical Education. 2010 Nov; 87(11):1123–1124. http://pubs.acs.org/doi/abs/10.1021/ed100697w, doi: 10.1021/ed100697w.

[45] NCI Open Database, August 2006 Release;. https://cactus.nci.nih.gov/download/nci/.

[46] Enhanced NCI Database Browser 2.2;. https://cactus.nci.nih.gov/ncidb2.2/.

[47] **Kim S**, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH. PubChem Substance and Compound databases. Nucleic Acids Research. 2016 Jan; 44(D1):D1202–D1213. https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkv951, doi: 10.1093/nar/gkv951.

[48] NCI/CADD Chemical Identifier Resolver;. https://cactus.nci.nih.gov/chemical/structure.

[49] **Bemis GW**, Murcko MA. The properties of known drugs. 1. Molecular frameworks. Journal of medicinal chemistry. 1996; 39(15):2887–2893.

[50] OEMedChem Toolkit 2017.Feb.1;. OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.

[51] Sirius T3 User Manual, v1.1. Sirius Analytical Instruments Ltd, East Sussex, UK; 2008.

[52] **Avdeef A**, Comer JEA, Thomson SJ. pH-Metric log P. 3. Glass electrode calibration in methanol-water, applied to pKa determination of water-insoluble substances. Analytical Chemistry. 1993; 65(1):42–49. https://doi.org/10.1021/ac00049a010, doi: 10.1021/ac00049a010.

[53] **Takács-Novák K**, Box KJ, Avdeef A. Potentiometric pKa determination of water-insoluble compounds: validation study in methanol/water mixtures. International Journal of Pharmaceutics. 1997; 151(2):235 – 248. http://www.sciencedirect.com/science/article/pii/S0378517397049077, doi: https://doi.org/10.1016/S0378-5173(97)04907-7.

[54] **Szakacs Z**, Beni S, Varga Z, Orfi L, Keri G, Noszal B. Acid-Base Profiling of Imatinib (Gleevec) and Its Fragments. Journal of Medicinal Chemistry. 2005; 48(1):249–255. https://doi.org/10.1021/jm049546c, doi: 10.1021/jm049546c, pMID: 15634018.

[55] **Szakacs Z**, Kraszni M, Noszal B. Determination of microscopic acid?base parameters from NMR?pH titrations. Analytical and Bioanalytical Chemistry. 2004 Mar; 378(6):1428–1448. http://link.springer.com/10.1007/s00216-003-2390-3, doi: 10.1007/s00216-003-2390-3.

[56] **Dozol H**, Blum-Held C, Guédat P, Maechling C, Lanners S, Schlewer G, Spiess B. Inframolecular acid–base studies of the tris and tetrakis myo-inositol phosphates including the 1, 2, 3-trisphosphate motif. Journal of molecular structure. 2002; 643(1-3):171–181.

[57] OEDepict Toolkit 2017.Feb.1;. OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.

[58] **Fraczkiewicz R**. In Silico Prediction of Ionization. In: *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering* Elsevier; 2013.http://linkinghub.elsevier.com/retrieve/pii/B978012409547202610X, doi: 10.1016/B978-0-12-409547-2.02610-X.